Cellular responses to oncogenic Ras signalling

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy

by

Craig James Mageean

September 2014

I dedicate this thesis to my father,

Alan Mageean,

who passed away in September 2013

following a hard-fought battle against cancer.

Dad, this is for you.

I - TABLE OF CONTENTS

II – FIGURE AND TABLE LIST

III – ACKNOWLEDGEMENTS

I thank my primary supervisor Professor Ian Prior for tolerating my presence in his lab for the past few years and, although we had our run-ins, I feel this experience overall has been a positive one and will always appreciate his understanding. I also thank my secondary supervisor Professor Alexey Tepikin, as well as Professors Chris Sanderson and Andrea Varro, whose actions and support helped at a time when I really needed it. I owe a huge debt of gratitude to Dr. Dean Hammond, without his backing and questionable humour, mostly at my expense, I doubt I would have made it to the end of this PhD. I also thank Dr. Hammond for aiding with the experiments and data analysis presented in the third part of this thesis. A big thanks goes to the now Dr. Ewan MacDonald for the countless pints and chats about football, as well as Drs. Emma Rusilowicz and Fiona Hood for forcing me out to the pub and lending a sympathetic ear. For their great support, I'm grateful to Dr. Simon Oliver, Alison Beckett and Dr. Maria Hernandez-Valladares. Maria, I share your views entirely. I'm also greatly indebted to our collaborators at the Manchester Institute, Dr. John Griffiths, Dr. Duncan Smith, Navin Chicooree and Yvonne Connolly, who provided a vibrant work environment and stimulated my passion for mass spectrometry; it was pleasure working with all of you. I also thank Liam, Veronica, Anna, Aitor, Amos, Arnaud, Becca and the many other lab members past and present. I also thank my old boss at Boehringer Ingelheim, Dr. Frank Büttner, for his advice and endless encouragement. I hugely appreciative the funding provided by the Wellcome Trust and for their agreement to generously fund a 6-month extension. To my mum and brother, we have been through some seriously tough times together of late and you two really held me up; I would never have reached this stage without you both, thank you. To my late father, you will forever be a limitless source of motivation and inspiration for me.

IV – ABSTRACT

The three human Ras proto-oncogenes encode for four highly homologous protein isoforms, H-, K(A)-, K(B)- and N-Ras, that function as molecular switches and timers to transduce signals from activated cell surface receptors to modulate signalling pathways responsible for cell proliferation, growth, survival and apoptosis. Oncogenic mutations of Ras occur in ~16% of all human cancer cases and most often affect codons 12, 13 and 61, each resulting in constitutive activation of the protein. Currently, it is uncertain how many Ras molecules are present per cell and which isoform is most abundant, essential information for the emerging field of systems biology, which utilises mathematical modelling to understand the behaviour of complex signalling pathways in a holistic approach. In the first part of this thesis, a selected reaction monitoring (SRM)-based Ras quantification technique is established for the measurement of each major isoform, along with the generation of isotope-labelled Ras protein standards that support quantification using the protein standard absolute quantification (PSAQ) strategy. Combining targeted SRM-based proteomics with the PSAQ strategy enabled the most accurate measurement of cellular Ras isoform abundance to date, as detailed in the second part of this thesis. The target cell lines in the present thesis were a panel of isogenic SW48 colorectal cells, which express a variety of heterozygous Ras mutations following the exchange of a wild-type allele for a mutant Ras sequence through targeted homologous recombination. Over 250,000 Ras molecules per wild-type SW48 cell were quantified, with K(B)-Ras the major Ras isoform. ~114 Ras molecules are predicted to be bound per μm^2 of plasma membrane, with the cellular molarity of Ras at 253 mM. Intriguingly, the presence of certain K-Ras mutations induced significant changes in cellular Ras abundance, which may be attributable to the specific transforming potential of each mutant protein. The presented data also suggests that, in addition to their structural and biochemical variations, Ras mutant proteins may exert their different biological effects through differential protein expression.

Many cancer types demonstrate a preference for a single Ras isoform to be mutated and present a codon-specific mutation signature. In colorectal cancer, K-Ras is the most frequently mutated isoform, with over 75% of mutations accounted for by G12D, G12V and G13D in colon tumours with a K-Ras mutation. Clinical data

suggests that patients harbouring codon 12 or 13 K-Ras mutations have different overall survival rates and responsiveness to treatment. Isogenic SW48 cells harbouring the aforementioned K-Ras mutations were subject to mass spectrometry-based proteome and phosphoproteome quantitative analysis, to investigate the effects of different amino acid substitution (G12D vs. G12V) or codon mutation (G12D vs. G13D) on oncogenic Ras signalling. The presented data in the third part of this thesis provide the first quantitative proteomic and phosphoproteomic profile of signatures associated with specific K-Ras mutant proteins expressed at endogenous levels. Each K-Ras mutation induced a distinct signalling output, with the most variability observed between codon 12 mutants and G13D. This indicates that the position of a Ras point mutation may have more impact on signalling than an amino acid substitution. Several proteins relevant to colorectal carcinogenesis were found to specifically up-regulated by codon 12 or 13 mutations, with one notable example being the recently described colon cancer stem cell marker double cortin-like kinase 1, which was sensitive only to codon 12 K-Ras mutations. Together, the data presented in this thesis provides a novel insight into the behaviour of a range of Ras mutations in a colorectal cancer setting.

V – ABBREVIATIONS

ACN	Acetonitrile
Ambic	Ammonium bicarbonate
AQUA	Absolute quantification
CE	Collision energy
CID	Collision-induced dissociation
CRC	Colorectal cancer
Da	Dalton
DDA	Data-dependent acquisition
dFBS	Dialysed foetal bovine serum
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
e.g.	Exempli gratia
ECD	Electron capture dissociation
ER	Endoplasmic reticulum
ESI	Electrospray ionisation
ETD	Electron transfer dissociation
eV	Electron Volt
FASP	Filter-aided sample preparation
FBS	Foetal bovine serum
FDA	False discovery rate
g	Gravity
GAP	GTPase-activating Protein
GDP	Guanosine diphosphate
GEF	Guanine nucleotide exchange factor
GO	Gene ontology
GTP	Guanosine triphosphate
i.e.	Id est
IAA	Iodoacetamide
iBAQ	Intensity-based absolute quantification
IDA	Information-dependent acquisition
iTRAQ	Isobaric tags for relative and absolute quantitation
LB	Luria Bertani agar/ broth
LC-MS	Liquid chromatography mass spectrometry
LMCO	Low mass cut off
LTQ	Linear trap quadrupole
m/z	Mass-to-charge ratio
MALDI	Matrix assisted laser desorption/ ionisation
MAPK	Mitogen-activated protein kinase
MIDAS	MRM-initiated detection and sequencing
min	Minute(s)
mRNA	Messenger ribonucleic acid

MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MSA	Multistage activation
nESI	Nanoelectrospray ionisation
nLC	Nanoflow liquid chromatography
PBS	Phosphate-buffered saline
PCA	Principle component analysis
PCR	Polymerase chain reaction
PI3K	Phosphatidylinositol-4,5-bisphosphate 3-kinase
ppm	Parts per million
PSAQ	Protein standard absolute quantification
PTM	Post-translational modification
QMF	Quadrupole mass filter/ transmission quadrupole
QqQ	Triple quadrupole mass spectrometer
R2	Squared Pearson correlation coefficient/ coefficient of determination
Ras	Rat sarcoma
rcf	Relative centrifugal force
RNA	Ribonucleic acid
RP-HPLC	Reversed-phase high performance liquid chromatography
rpm	Revolutions per minute
RT	Room temperature
SCX	Strong cation exchange
SD	Standard deviation
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SILAC	Stable isotope labelling of amino acids in cell culture
SRM	Selected-reaction monitoring
TBS	Tris-buffered saline
TCA	Tricarboxylic acid
TFA	Trifluoroacetic acid
TIC	Total ion current
V	Volt
XIC	Extracted ion chromatogram

1 – INTRODUCTION

1.1 RAS GENES, ISOFORMS, STRUCTURE AND SIGNALLING

Ras small GTPases have become model proteins for signal transduction and molecular oncology research. Over the past 30 years, they have played a significant role in understanding the origins of cancer at the molecular level, establishing fundamental concepts in signal transduction and initiating the discovery of the large Ras protein superfamily (Cox and Der, 2010; Malumbres and Barbacid, 2003). Despite these research efforts, a comprehensive understanding of Ras function and dysfunction, particularly in human cancer, is yet to come. Here the work that led to the discovery of Ras and its characterisation as an oncogene is discussed, along with the current knowledge of Ras structure, isoforms, splice variants, expression levels and signalling.

1.1.1 THE DISCOVERY OF RAS

Ras was discovered as a result of extensive research of acutely transforming retroviruses isolated from mice, rats, cats and other animals; a field that dates back to the early 1900s with Vilhelm Ellerman, Olaf Bang and Peyton Rous' discoveries of transmissible leukaemia and solid tumours in chickens (Ellermann and Bang, 1909; Rous, 1910). However, it was in 1964 and 1967 when we had our first glimpses of the genetic elements that we now recognise as HRAS and KRAS, respectively. In 1964, Jennifer Harvey found that a preparation of Maloney's leukaemogenic virus, taken from a leukaemic rat, induced sarcomas when injected into new-born mice (Harvey, 1964). Three years later, Kirsten and Mayer discovered that new-born rats developed multiple sarcomas following injection with spleen filtrates or plasma from rats infected with mouse erythroblastosis virus (Kirsten and Mayer, 1967). The Ras gene name was coined in the 1970s, after a new nomenclature for retroviral oncogenes was proposed, based on three-letter acronyms. Since Harvey and Kirsten oncoviruses cause rat sarcomas it became the basis of the *ras* gene name, with the discoverers' names used to distinguish between them (Malumbres and Barbacid, 2003). During this time, Scolnick and colleagues suggested that incorporation of normal rat sequences into the viral genome was responsible for the transforming properties of the Harvey and Kirsten oncoviruses (Scolnick and Parks,

1974; Scolnick et al., 1973). However, molecular characterisation of these sequences was held back until restrictions on recombinant DNA technology, imposed during the Asilomar conference in 1975, expired. In the late 1970s and early 1980s, Ras research progressed at an astonishing pace, with Scolnick and colleagues delivering a series of fundamental discoveries that established the basic biochemical and cellular properties of Ras. They determined Harvey and Kirsten Ras genes encode 21 kDa proteins (Parks and Scolnick, 1977; Shih et al., 1979b), which bind GDP and GTP (Scolnick et al., 1979) and associate with the plasma membrane (Willingham et al., 1980). Moreover, they also detected Ras protein in human cells and those from various vertebrate species using a rat-derived anti-Ras antibody (Langbeheim et al., 1980) and identified Ras genes in rat, human and other vertebrate genomes (Chang et al., 1982b; DeFeo et al., 1981; Ellis et al., 1981). Also in the early 1980s, DNA transfection assays, using the transformation of mouse NIH-3T3 cells as readout, were heavily utilised to detect activated oncogenes in tumour-derived genomic DNA (Shih et al., 1981; Shih et al., 1979a). Several studies isolated oncogenes from the genomic DNA of various carcinomas, including human bladder, lung, colon and others; but the exact nature of the isolated genes was unknown (Goldfarb et al., 1982; Krontiris and Cooper, 1981; Murray et al., 1981; Perucho et al., 1981; Shih and Weinberg, 1982). The transforming genes were only exposed as cellular homologs of the viral Ras genes when transformed NIH-3T3 genomic DNA was hybridised with retroviral probes (Der et al., 1982; Parada et al., 1982; Santos et al., 1982). Soon after, it was demonstrated that overexpression of human H-Ras could transform NIH-3T3 cells (Chang et al., 1982a). Later in 1982, molecular cloning of normal and oncogenic H-Ras genes revealed the mechanism of Ras oncogene activation as a single point mutation at residue 12 or less commonly at 13 and 61 (Capon et al., 1983a; Reddy et al., 1982; Tabin et al., 1982; Taparowsky et al., 1982). This made H-Ras the first cell-derived oncogene whose mechanism of activation was elucidated at the molecular level. By 1983, Wigler and Weiss groups discovered the third member of the Ras gene family, designated N-Ras, by identifying transforming sequences from neuroblastoma cell lines (Hall et al., 1983; Shimizu et al., 1983). In 1984, the discovery of mutated K-Ras oncogenes in tumour biopsies, but not in the normal tissue of the same patients (Nakano et al., 1984; Santos et al., 1984), showed that Ras oncogene activation occurred in the wild and not only in the laboratory. The following section covers the biochemical and physiological properties of Ras, while section 2 describes the role of Ras in cancer.

1.1.2 RAS GTPases: PROTO-ONCOGENES, ISOFORMS AND SYNTHESIS

Ras proteins are signal-transduction GTPases that sit near the top of signalling cascades and coordinate a wide variety of cellular functions, including cell proliferation and survival. They are also proto-oncogenes that are frequently mutated in human cancer. Mammalian cells ubiquitously express the three oncogenic members of the Ras family: H, K and N, which are encoded by separate genes and share a high degree of sequence homology, but are not functionally redundant (Fiorucci and Hall, 1988; Furth et al., 1987; Karnoub and Weinberg, 2008; Prior et al., 2012). Ras GTPases act as molecular switches and timers, alternating between GTP- and GDP-bound states, which differ by the conformation of their nucleotide-sensitive switch 1 and 2 regions. The GDP-bound state is generally considered inactive, while the GTP-bound state is a prolific signalling molecule that can bind to over twenty different effectors to activate a network of downstream signalling events (Goldfinger et al., 2007; Prior and Hancock, 2012). Since Ras proteins tightly bind GDP and hydrolyse GTP very slowly, they require the helping hand of guanine nucleotide exchange factors (GEFs), which facilitate dissociation of GDP, and GTPase activating proteins (GAPs) that stimulate GTP hydrolysis (Bos et al., 2007). Thus, the actual molecular switch is comprised of a Ras protein, a GAP and a GEF, rather than Ras alone.

1.1.2A RAS GENES AND SPLICE VARIANTS

Human cells harbour 3 Ras proto-oncogenes that encode for H, K and N isoforms. The HRAS gene (11p15.5) is the smallest at 5,046 bases in size, then NRAS (1p13.2) at 12,431 bases, with the KRAS gene (12p12.1) the largest at 46,148 bases. The HRAS gene comprises of 4 exons with an alternative exon, denoted ‘intron D exon’ (IDX), located in the last intron (Cohen et al., 1989). A point mutation in the IDX sequence was shown to increase H-Ras levels ten-fold (Cohen and Levinson, 1988), suggesting this region was a negative-acting regulatory element. However, alternative splicing of H-Ras pre-mRNA was shown to render two mRNAs that encode for p21 H-Ras or p19 H-RasIDX that includes the IDX sequence (Cohen et al., 1989). The KRAS gene is composed of exons 1, 2, 3, 4A and 4B. Alternative splicing of the fourth exon renders two mRNAs that encode two different protein isoforms: K(A)- and K(B)-Ras (Capon et al., 1983b; McGrath et al., 1983). K(B)-Ras

is the predominant splice variant and ubiquitously expressed, whereas K(A)-Ras is not (Plowman et al., 2006). The functional differences between K(A)- and K(B)-Ras, and other Ras isoforms, are discussed in detail in section 1.3. Recently, alternative splicing of NRAS pre-mRNA has been reported to render 5 different mRNAs that encode 5 proteins of various lengths (20 to 208 amino acids) (Eisfeld et al., 2014).

1.1.2B RAS ISOFORMS AND THEIR STRUCTURE

The three mammalian Ras proto-oncogenes encode for 4 major isoforms, H-, K(A)-, K(B)- and N-Ras, which are 21 kDa proteins, 188/9 amino acids in length and share almost complete sequence homology between residues 1-165, a region known as the G-domain that encompasses the nucleotide and effector binding motifs (Fig. 1.1A). All major isoforms, bar K(A)-Ras, are ubiquitously expressed and are principally distinguished from each other by the C-terminal hypervariable region (HVR, final 23-24 amino acids) where there is < 15% sequence similarity (Fiorucci and Hall, 1988; Omerovic et al., 2007). Ras proteins, like all small GTPases, are comprised of a six-stranded β-sheet surrounded by five α-helices (de Vos et al., 1988; Pai et al., 1989). On GDP-GTP exchange, a 'conformational domino effect' propagates throughout the Ras molecule and causes two nucleotide-sensitive regions to 'switch' their conformation. These surface-localised regions are named switch 1 and 2. Switch 1 spans residues 30 to 38 (corresponding to most of loop 2) and switch 2 comprises of residues 59 to 76 (loop 4 and α-helix 2) (Fig. 1.1 B) (Milburn et al., 1990). These subtle, disorder-to-order structural transitions reveal an effector-binding site on the Ras molecule (see section 1.3.1 for further details) and occur due to the presence of GTP's Υ-phosphate in the Ras nucleotide-binding pocket (Fig. 1.2). The binding pocket is primarily composed of loops L1, L2, L4, L8 and L10; with guanine specificity conferred by DTAGQ-, SAK- and NKxDL-motifs. The aspartate side chains from the DTAGQ- and NKxDL-motifs (residues 57-61 and 116-120, respectively) form a bifurcated hydrogen bond with the guanine base, while alanine from the SAK-motif (residues 145-147) would not tolerate an adenine amino group due to steric hindrance (Vetter and Wittinghofer, 2001). As a result, Ras has a nanomolar to picomolar affinity for GDP/ GTP and millimolar affinity for ADP/ ATP (Scolnick et al., 1979).

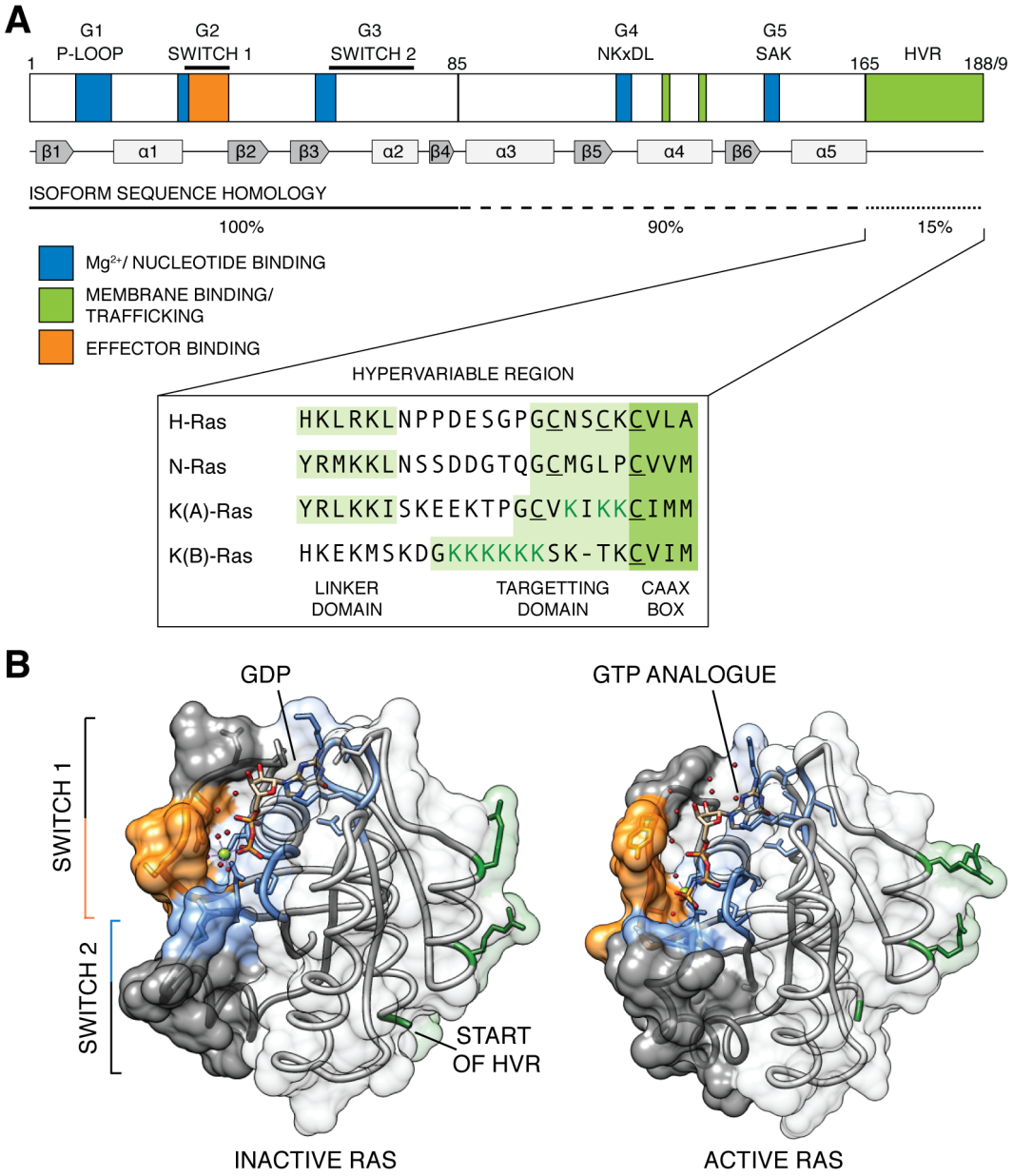


Fig. 1.1. The structure of Ras. (**A**) H-, K(A)-, K(B)- and N-Ras isoforms consist of a series of ten loops that interconnect six β-sheets and five α-helices, with the main sequence differences between isoforms located in the C-terminal hypervariable region (HVR). The HVR is post-translationally modified to enable membrane interaction and differential localisation. Key functional domains and motifs, common to all Ras family members, are indicated. Sequences and basic residues important for membrane affinity are highlighted in green; underlined CAAX-box cysteine residues are farnesylated following synthesis; remaining underlined cysteine residues are targeted for palmitoylation; linker domain has a role in nanoclustering. (**B**) Crystallographic structures of inactive H-Ras-GDP (residues 1-171, 2.0 Å resolution) and active H-Ras-GppNHp (5'-Guanylyl imidodiphosphate) (residues 1-166, 1.26 Å resolution) show the different conformations of the nucleotide-sensitive switch 1 (residues 30-38) and switch 2 (residues 60-76) regions when GDP or GTP is bound to Ras. Structures are coloured according to A; areas of reduced transparency indicate switch 1 and 2 regions; Mg^{2+} ion shown as a green ball; water molecules as red balls. H-Ras-GDP from (Milburn et al., 1990), Protein Data Bank (PDB) code 4q21; H-Ras-GppNHp from (Scheidig et al., 1999) , PDB code 1ctq.

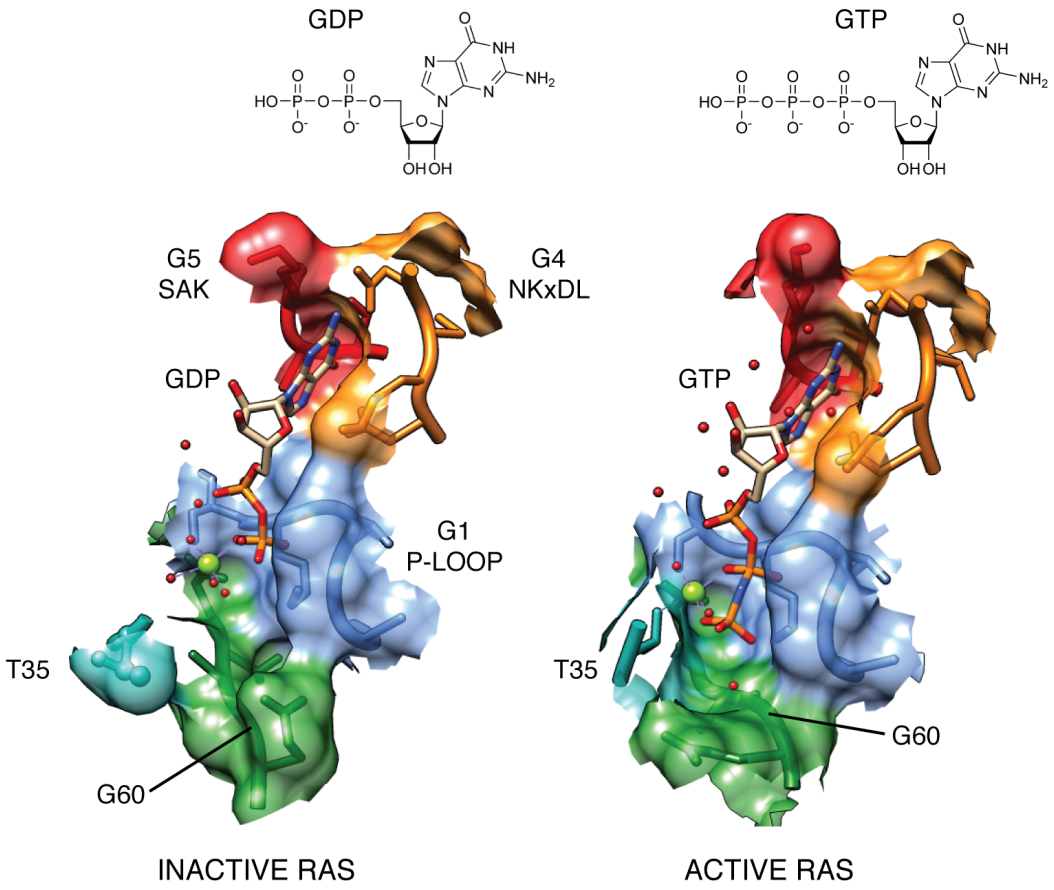


Fig. 1.2. The guanine nucleotide-binding pocket of Ras. (Left) Binding site of inactive H-Ras-GDP and (right) of active H-Ras-GppNHp (5'-Guanylyl imidodiphosphate). Five G-motifs comprise the nucleotide-binding pocket; the P-loop (blue) responsible for binding with the β-phosphate of GDP or GTP; T35 of switch 1 (turquoise) binds to the ϒ-phosphate of GTP and a Mg^{2+} ion (green ball) but makes no contact with GDP; DTAGQ-motif of switch 2 (green), D57 confers specificity for guanosine over adenosine, Q61 is critical for alignment of a water molecule (red balls) for GTP hydrolysis; NKxDL- (orange) and SAK-motifs (red) provide further selectivity for guanosine. The ϒ-phosphate of GTP induces structural rearrangements of nucleotide-sensitive switch 1 and 2 regions, through interactions with T35 and G60. Inactive Ras binding-site data from (Milburn et al., 1990), Protein Data Bank (PDB) code 4q21; Active Ras from (Scheidig et al., 1999), PDB code 1ctq.

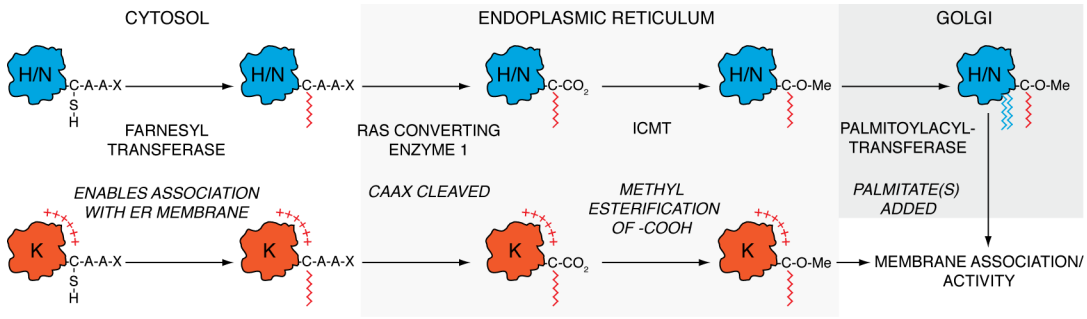


Fig. 1.3. Synthesis and processing of Ras. Ras proteins are synthesised in the cytosol by free ribosomes, where farnesyl transferase recognises their CAAX motif. Farnesylation confers some membrane binding ability to Ras, enabling it to associate with the ER membrane. At the ER, Ras encounters Ras converting enzyme, which removes the CAAX motif, before isoprenylcysteine carboxyl methyltransferase (Icmt) methyl esterifies the subsequent carboxyl group. K(B)-Ras, with its polybasic domain (red pluses), now has the necessary membrane-binding capability to associate with lipid membranes. H- and N-Ras require palmitoylation by Golgi-resident Ras palmitoyltransferase to enable plasma membrane binding. These isoforms then enter the exocytic pathway on their way to the plasma membrane. K(A)-Ras follows the trafficking route as of H- and N-Ras.

The phosphate-binding loop (P-loop, residues 10-17) creates a polyanion hole that binds to and neutralises the negative charge of the nucleotide's β-phosphate through residues 13-17 (Pai et al., 1989; Pai et al., 1990; Saraste et al., 1990; Valencia et al., 1991). Ras structural rearrangement is triggered when the γ-phosphate of GTP hydrogen bonds with the main chain amino groups of switch 1's T35 and switch 2's G60 (part of the DTAQ-motif) (Milburn et al., 1990; Vetter and Wittinghofer, 2001). Upon GTP hydrolysis, the residues within switch 1 and 2 revert back to their inactive states. A helpful analogy for Ras conformational change is that of a loaded spring mechanism, where release of the γ-phosphate following GTP hydrolysis permits the two switch regions to relax back to their GDP-bound state.

1.1.2C RAS SYNTHESIS AND PROCESSING

Ras proteins are synthesised in the cytosol as globular hydrophilic precursors by polysomes and need subsequent post-translational modification to enable them to stably associate with membranes, an essential requirement for biological activity (Willumsen et al., 1984). The initial triplet of modifications is directed against the C-terminal CAAX-motif (in which A is any aliphatic amino acid and X any amino acid) that is common to all major Ras isoforms (Fig. 1.1A and 1.3). Firstly, the cysteine residue is farnesylated by farnesyl transferase (Reiss et al., 1990), which facilitates association to the cytosolic surface of the endoplasmic reticulum (ER). Secondly, an ER-localised endopeptidase, Ras and a-factor convering enzyme (Rce1), cleaves

the AAX tripeptide (Boyartchuk et al., 1997; Gutierrez et al., 1989; Kim et al., 1999). Thirdly, the C-terminal farnesyl-cysteine is methylated by isoprenylcysteine carboxyl methyltransferase (Icmt) (Dai et al., 1998; Hrycyna et al., 1991). Interestingly, K(B)-Ras appears to be more efficiently methylated than H- or N-Ras, but it unclear why this is the case (Choy et al., 1999; Hancock, 2003). Following methylation, Ras proteins diverge on their path to the plasma membrane. H-, K(A)- and N-Ras are palmitoylated once or twice on cysteine residues in their HVRs by Ras palmitoyltransferase (DHHC9 in mammals), before entering the exocytic pathway and trafficking through the Golgi to their final cell membrane destinations (Apolloni et al., 2000; Casey et al., 1989; Choy et al., 1999; Swarthout et al., 2005) (Fig. 1.1A and 3). K(A)- and N-Ras, in addition to their mono-palmitoylation, also utilise hydrophobic/ basic residues in their linker domains for stable membrane localisation (Laude and Prior, 2008). The polybasic, lysine-rich region of K(B)-Ras enables it to electrostatically interact with negatively charged phospholipid head groups to support membrane association (Silvius et al., 2006). It also allows K(B)-Ras to forgo palmitoylation and bypass the Golgi on its way to the plasma membrane (Hancock et al., 1990), possibly through interactions with microtubules or with class C Vps proteins as observed in yeast (Chen et al., 2000; Thissen et al., 1997; Wang and Deschenes, 2006). Phosphodiesterase 6δ (PDEδ) may also have a role in chaperoning K(B)-Ras during intracellular trafficking, by shielding its hydrophobic farnesyl group from the aqueous environment of the cytosol (Chandra et al., 2012). However, the exact mechanisms underlying K(B)-Ras trafficking remain unclear.

1.1.3 RAS SIGNALLING: EFFECTOR BINDING, REGULATION, ISOFORM- AND COMPARTMENTALISED-SIGNALLING

Ras GTPases activate signalling pathways by binding to effectors and recruiting them to membranes. This can cause conformational change in the effectors and/ or facilitate interactions with activating proteins/ co-factors necessary for signal propagation (Prior and Hancock, 2012). Ras activation can occur following ligand binding with receptor tyrosine kinases, e.g. epidermal growth factor receptor (EGFR), that autophosphorylate before binding adapter proteins, e.g. GRB2, which recruit GEFs to the plasma membrane to facilitate Ras GDP-GTP exchange. Ras-GTP then transduces the extracellular stimulus into a network of intracellular signalling events (Fig. 1.4). The plasma membrane is a critical platform for Ras

signalling, but endomembranes provide additional stages from which Ras can signal. Here various aspects of Ras signalling are discussed, including Ras effector binding, regulation of Ras activity with GEFs/ GAPs, isoform-specific and compartmentalised Ras signalling and nanoclustering.

1.1.3A RAS EFFECTOR REGION AND RAS INTERACTING DOMAINS

Ras proteins primarily bind with effectors through residues 32-41, which overlap with the switch 1 region (Sigal et al., 1986). The Ser/Thr protein kinase c-Raf1 binds to switch 1 through a stable 81-residue domain, known as the Ras binding (RB) domain (Fig. 1.5) (Koide et al., 1993; Moodie et al., 1993; Nassar et al., 1995). The RB domain has an ubiquitin fold structure and interacts with Ras through complementary charge interactions and the formation of an inter-protein β-sheet (Fig. 1.5) (Emerson et al., 1995; Nassar et al., 1996). Other Ras effectors, e.g. RalGDS (Ral guanine dissociation stimulator), do not contain an RB domain but a Ras association (RA) domain (Ponting and Benjamin, 1996). Despite low sequence homology between RB and RA domains, they both have similar structures and bind to switch 1 of Ras in a similar manner (Geyer et al., 1997). Interestingly, RA domains can also interact with switch 2 of Ras (Huang et al., 1998). This involvement of switch 2 in protein-protein interactions is not an isolated case, as phosphoinositide 3-kinase (PI3K) gamma interacts with both Ras switch regions using its RB and catalytic domains (Fig. 1.5B) (Pacold et al., 2000).

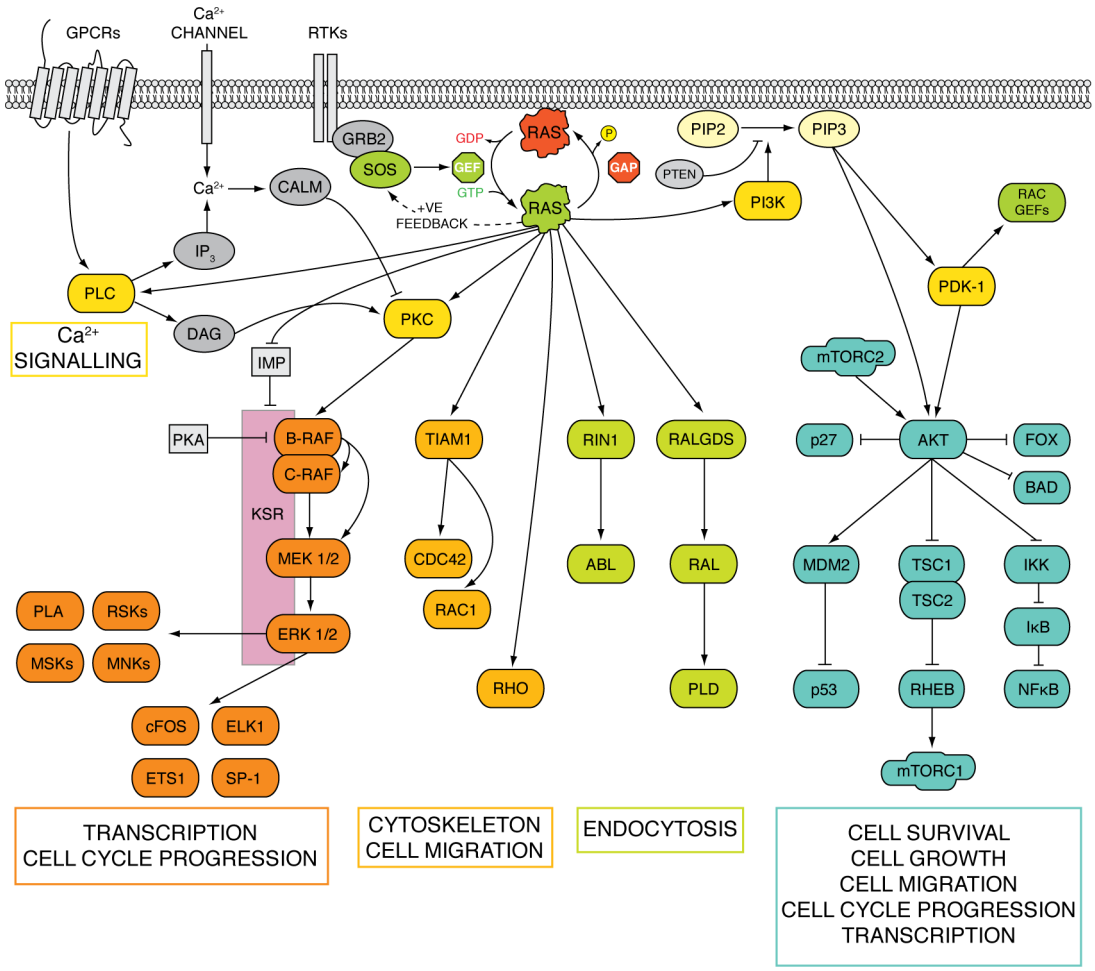


Fig. 1.4. Overview of major Ras signalling pathways. Active (farnesylated, membrane- and GTP-bound) Ras stimulates a network of downstream signalling cascades, with the mitogen-activated protein kinase (MAPK, dark orange) and phosphoinositide 3-kinase (PI3K)-Akt (turquoise) pathways the two best characterised. The MAPK pathway is an essential element of mitogenic signalling involving receptor tyrosine kinases (RTKs) that trigger a variety of cellular processes crucial for cell growth, differentiation and apoptosis (Dhillon et al., 2007). Active Ras binds to the Raf family of Ser/ Thr kinases (Raf-1, A-Raf and B-Raf) to recruit them to membranes, where they are activated through phosphorylation by various kinases (Marais et al., 1998; Marais et al., 1995). Active Raf then phosphorylates MEK, which in turn activates extracellular signal regulated kinases 1 and 2 (ERK1/2). Active ERK1/2 then phosphorylates hundreds of cytoplasmic and nuclear substrates, including regulatory molecules and transcription factors (Roskoski, 2012). Scaffold proteins play an essential role in spatiotemporal regulation of MAPK signalling. Kinase suppressor of Ras (KSR) is one of several MAPK scaffolds and is constitutively associated with MEK1/2, only becoming associated with ERK1/2 following cellular stimulation (Brennan et al., 2011; Claperon and Therrien, 2007). The second best-characterised Ras effector family is the PI3K family, which have essential function on cell survival and growth (Castellano and Downward, 2010). Active PI3K converts phosphatidylinositol (4,5)-bisphosphate (PIP_2) into phosphatidylinositol (3,4,5)-trisphosphate (PIP_3). PIP_3 binds to the plekstrin homology domain of AKT stimulating its kinase activity, leading to phosphorylation of a host of proteins that coordinate cell cycle entry, cell growth and survival. Another well-characterized Ras effector is Ral guanine nucleotide dissociation stimulator (RalGDS) (Ferro and Trabalzini, 2010). These proteins link Ras signaling with activation of RalA and RalB small GTPases.

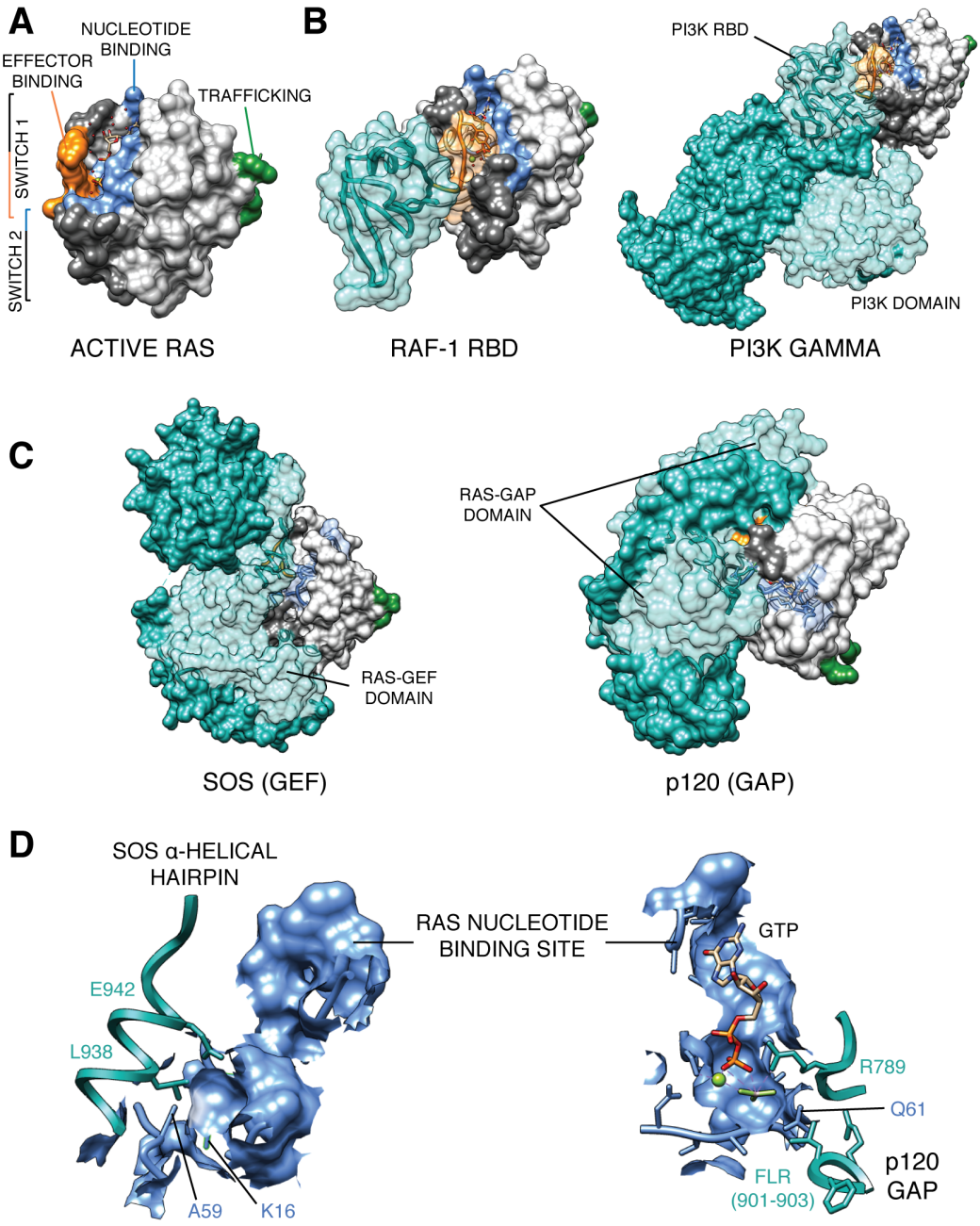


Fig. 1.5. Ras effector/ regulator complexes. (**A**) Active H-Ras-GppNHp (PDB code 1ctq). (**B**) Ras binding with its effectors. Left, Structure of H-Ras-GppNHp complexed with the Ras binding domain (RBD) of Raf-1 kinase (2.5 Å resolution, PDB code 4g0n). Right, H-RasG12V in complex with phosphoinositide 3-kinase (PI3K) gamma catalytic subunit (3.0 Å resolution, PDB code 1he8; (Pacold et al., 2000)). Effectors bind with Ras using an interface within the switch 1 region (residues 32-41), but Ras also uses its switch 2 region to bind PI3K. (**C**) Ras binding with its regulators. Left, H-Ras-GppNHp in complex with the plekstrin homology domain of Son of Sevenless (SOS) (2.8 Å resolution, PDB code 1bkd; (Boriack-Sjodin et al., 1998)). Right, H-Ras in complex with p120 Ras GTPase-activating protein (p120GAP) (2.5 Å resolution, PDB code 1wq1; (Scheffzek et al., 1997)). (**D**) Left, Guanine nucleotide exchange (GEF) factor function. To eject GDP from the Ras binding pocket, SOS inserts its α-helical hairpin that reorientates A59 and disturbs K16 binding with GDP. Right, GAP function. p120GAP inserts a catalytic R789 finger into the Ras nucleotide binding site that reorientates Q61, which aligns a water molecule to enhance Ras GTPase activity whilst stabilising the protein.

1.1.3B GUANINE EXCHANGE FACTORS AND GTPase ACTIVATING PROTEINS

The GDP-GTP cycle of Ras is highly regulated by the coordinated action of GEFs and GAPs, of which there are 27 members of each in human cells. Interestingly, the Ras family, which has 36 members, outnumber their regulators (Cherfils and Zeghouf, 2013). As mentioned earlier, Ras is an incomplete enzyme that binds guanine nucleotides with high affinity and therefore must enlist the help of GEFs and GAPs to complete the molecular switch. The catalytic domains of many Ras GEFs are structurally unrelated and approach the G-proteins from different angles, yet they all utilise similar principles to deform the nucleotide-binding pocket, particularly the phosphate-binding region, to facilitate GDP ejection. For example, SOS makes extensive contacts with switch 2 and uses a α-helical hairpin to force open the binding site (Fig. 1.5C and D) (Boriack-Sjodin et al., 1998). GEFs sterically occlude the magnesium-binding site or reposition the alanine side-chain from the DTAG-motif of switch 2, which in turn disturbs P-loop binding with the GDP's β-phosphate (Bos et al., 2007).

Ras GAPs can accelerate the intrinsic GTPase activity of Ras by five orders of magnitude (Gideon et al., 1992), through interactions that optimally align a water molecule for nucleophilic attack on GTP's γ-phosphate (Fig. 1.5 C and D) (Scheffzek et al., 1997). For example, p120 GAP inserts a catalytic arginine finger into the nucleotide-binding pocket that alters the position of Q61, which in turn optimally aligns the attacking water molecule. The arginine finger also stabilises the transition state by neutralising the negative charge of the γ-phosphate (Bos et al., 2007).

For some small GTPases, mainly Rab and Rho proteins, guanine-dissociation inhibitors (GDIs) provide an additional level of control. These proteins induce small GTPases to dissociate from membranes by sequestering prenyl groups on their lipidated tails and maintain them in the GDP-bound state (Bos et al., 2007; Cherfils and Zeghouf, 2013). So far, only PDEδ has been described as a GDI-like solubilising factor that interacts with farnesylated Ras proteins (Chandra et al., 2012).

1.1.3C ISOFORM-SPECIFIC RAS SIGNALLING

Ras isoforms were originally considered to be functionally identical because they exhibit a high degree of sequence homology, appeared to be governed by a common cohort of regulators and activate the same group of effectors (Bos, 1989; Castellano and Santos, 2011; Chesa et al., 1987). However, the dramatic asymmetry of Ras isoform mutations in certain cancer types, e.g. the frequent mutation of K-Ras in colon, pancreatic and lung cancers (discussed further in section 2) (Bos, 1989; Bos et al., 1987), provided the first clues that H-, K- and N-Ras had non-redundant roles.

Although H-, K(A)-, K(B)- and N-Ras qualitatively activate the same effector pathways, they have quantitative differences in their ability to activate particular downstream effectors. It is apparent that Ras isoforms differ in their ability to activate Raf-1, Rac and phosphoinositide 3-kinase (PI3K), with K(B)-Ras the most potent activator of Raf-1 and Rac, while H-Ras is an efficient activator of PI3K (Voice et al., 1999; Walsh and Bar-Sagi, 2001; Yan et al., 1998). K(A)-Ras activates Raf-1 more strongly than H/ N-Ras, but not K(B)-Ras, and along with N-Ras enabled anchorage-independent growth of RIE-1 cells (Voice et al., 1999) H- and K(B)-Ras also induce higher activation of NFκB than N-Ras (Millan et al., 2003). Such differential activation of effectors has also been demonstrated for other members of the Ras protein superfamily, e.g. R-Ras, TC21, Rheb, Rin and Rap 1 and 2 (Rodriguez-Viciana et al., 2004). In addition, H- and N-Ras may regulate the mitogen activated protein kinase (MAPK) pathway through different mechanisms, as treatment with a growth factor antagonist blocked proliferation of H-Ras transformed cells, but not those transformed with N-Ras (Hamilton and Wolfman, 1998). Furthermore, overexpression of Ras isoforms in isogenic colorectal cell lines revealed K-Ras is easily the most potent inducer of a transformed phenotype (Keller et al., 2007a).

Ras isoforms also differ in their sensitivity to GAPs and GEFs. Regarding GEFs, Ras-GRF induces potent *in vivo* GDP-GTP exchange on H-Ras but to a lesser extent on K- and N-Ras (Jones and Jackson, 1998; Matallanas et al., 2003). RasGRP2 can activate K- and N-Ras but not H-Ras (Clyde-Smith et al., 2000), while RasGRP3 demonstrates the opposite behaviour (Lorenzo et al., 2001; Ohba et al.,

2000). Furthermore, the ubiquitously expressed GEF Sons of Sevenless (SOS) can facilitate nucleotide exchange in all Ras isoforms, but demonstrates different degrees of potency: H > N > K (Jaumot et al., 2002). Considering GAPs, NF-1 and p120-GAP induce GTP hydrolysis at similar rates in H- and N-Ras, but N-Ras has slightly higher hydrolysis rates (Bollag and McCormick, 1991).

Genetic studies in mice showed that the K-Ras gene is essential for embryo development (Johnson et al., 1997; Koera et al., 1997), while H- and N-Ras genes are dispensable (Esteban et al., 2001; Umanoff et al., 1995). K(A)-Ras is also unessential for correct development (Plowman et al., 2003). This indicates that the K-Ras gene has unique function during embryogenesis that H- and N-Ras genes cannot compensate for. However, if the K-Ras gene is replaced with H-Ras it supports normal mouse embryonic development, but when the mice reach adulthood they display cardiomyopathy and hypertension (Potenza et al., 2005). Thus, it could be that Ras expression levels is a critical survival factor during mouse development and that H-Ras, when expressed under the K-Ras promoter, can substitute for K-Ras in its essential functions. Yet, it also suggests that K-Ras has a role in cardiovascular homeostasis that H-Ras cannot replicate.

1.1.3D COMPARTMENTALISED RAS SIGNALLING

Micro- and macro-localisation of Ras proteins at the plasma membrane and Golgi, ER, endosomal and mitochondrial endomembranes determine Ras signal output (Philips, 2005; Prior and Hancock, 2012). Although the cell surface represents the prime location for all isoforms, Ras proteins also signal from endomembranes. In most cell types, the relative contribution to the endomembranous Ras component is typically H ≥ N, K(A) > K(B) (Omerovic and Prior, 2009). Since Ras effectors, regulators and scaffolding proteins also display differential localisations throughout the cell, signalling from endomembranes is likely to engage different signalling pathways and therefore exert alternate biological outputs compared to the plasma membrane (Fig. 1.6A and B).

The importance of Ras cellular localisation can be demonstrated in fission yeast, which have a single Ras protein: Ras1. By Restricting Ras1 to endomembranes morphogenesis was triggered, whereas plasma membrane-restricted Ras1

stimulated the mating response (Onken et al., 2006). The ER and Golgi have provided some of the best evidence in support of compartmentalised Ras signalling. Unlike the transient activation of Ras at the plasma membrane, Ras signalling at the ER/ Golgi is sustained and can be activated through a variety of growth factors (Chiu et al., 2002). Golgi/ ER-resident Ras activation can occur through *in situ* GEF activation, e.g. Golgi-RasGRP1 is activated following PLC-ϒ generation of second messengers (Bivona et al., 2003), or diffusion of active Ras from the plasma membrane (Fehrenbacher et al., 2009). Interestingly, tethering of oncogenic H-RasG12V to different organelles showed that only Golgi-localised Ras was unable to induce cellular transformation (Matallanas et al., 2006). In contrast, Golgi-targeted Raf more potently activated ERK compared to ER-targeted Raf (Inder et al., 2008).

K(B)-Ras demonstrates almost complete plasma membrane localisation, most likely because its polybasic domain is preferentially attracted to the plasma membrane's net negative charge versus other intracellular locations, and exhibits membrane interactions with very short half lives (~ 7 mins) (Silvius et al., 2006; Yeung et al., 2006). This short membrane residency means that K(B)-Ras cannot traffic between compartments via membrane carriers. Instead, K(B)-Ras is phosphorylated in its HVR at S181 by protein kinase C (PKC) that triggers translocation to the cytosol and accumulation on the outer mitochondrial, ER and Golgi membranes (Fig. 1.6A) (Ballester et al., 1987; Bivona et al., 2006; Prior and Hancock, 2012). The addition of a phosphate group acts as an electrostatic switch, as its strong negative charge partially neutralises the hexalysine sequence causing K(B)-Ras plasma membrane dissociation. K(B)-Ras may also dissociate from the plasma membrane due to interactions with calmodulin that binds with the HVR of K(B)-Ras. This triggers reversible localisation to the Golgi and early/ recycling endosomes, providing a mechanism for Ca^{2+} signalling to coordinate K(B)-Ras localisation (Fivaz and Meyer, 2005; Villalonga et al., 2001). Intriguingly, there seems to be an interplay between PKC phosphorylation and calmodulin in K(B)-Ras regulation, as calmodulin can prevent S181 phosphorylation (Alvarez-Moya et al., 2011).

The localisation of H- and N-Ras is mainly determined by the presence and positioning of palmitoyl groups. Palmitoylation, which occurs only at the Golgi, promotes stable tethering of otherwise cytosolic proteins to intracellular membranes and is the only reversible lipid modification (Smotrys and Linder, 2004). Mutagenesis

of H-Ras C181 and C184 revealed that mono-palmitoylation of C184 traps H-Ras on the Golgi, whereas mono-palmitoylated C181 H-Ras localises to the plasma membrane (Roy et al., 2005). Moreover, H- and N-Ras proteins undergo cycles of acylation and deacylation that enable rapid trafficking between the plasma membrane and Golgi via a non-vesicular pathway (Fig. 1.6A) (Goodwin et al., 2005; Rocks et al., 2005). Since Ras half-life is around 20 hours (Ulsh and Shih, 1984) and that of palmitoylation is estimated to between a few minutes and hours, H- and N-Ras proteins are likely to go through many depalmitoylation/ repalmitoylation cycles during their existence (Baker et al., 2003; Magee et al., 1987). H- and N-Ras can also localise to the early/ recycling endosomes following growth factor stimulation or by di-ubiquitylation (Roy et al., 2002). The 19 kDa variant of H-Ras, H-RasIDX, exhibits a nuclear/ cytosolic localisation as it lacks the H-Ras HVR, due to an in-frame stop codon in the IDX sequence, meaning it cannot associate with membranes (Cohen et al., 1989; Guil et al., 2003). Interestingly, it does not interact with known Ras effectors Raf-1 or Rin-1, but interacts with RACK1 and p73β, which could indicate a role in organising scaffolding proteins and apoptosis/ cell cycle arrest (Guil et al., 2003; Jeong et al., 2006). Furthermore, cytosolic Ras has been shown to have a more pronounced effect on *Drosophila* eye development than membrane-bound Ras (Sung et al., 2010).

1.1.3E RAS PLASMA MEMBRANE NANOCLUSTERS

The plasma membrane is a complex, dynamic, non-equilibrium mix of over 7,000 different phospholipids, ~30-40 mol% cholesterol and 25% integral and peripheral membrane proteins; beneath which is an asymmetrical actin mesh that impedes lateral movement of transmembrane proteins (Prior and Hancock, 2012). This results in non-random distribution of lipids and proteins throughout the membrane.

Upon plasma membrane binding, Ras isoforms rapidly diffuse and segregate into distinct, non-overlapping nanoclusters (Fig 1.6A) (Abankwa et al., 2007). Each cluster is estimated to be 12-24 nm in diameter, to comprise of 6-7 Ras proteins and have a half-life of 0.1-1 sec (Hancock and Parton, 2005; Plowman et al., 2005). Ras nanoclusters form due to the interplay between the C-terminal lipid anchor, adjacent hypervariable regions and the G-domain with certain phospholipids/ cholesterol.

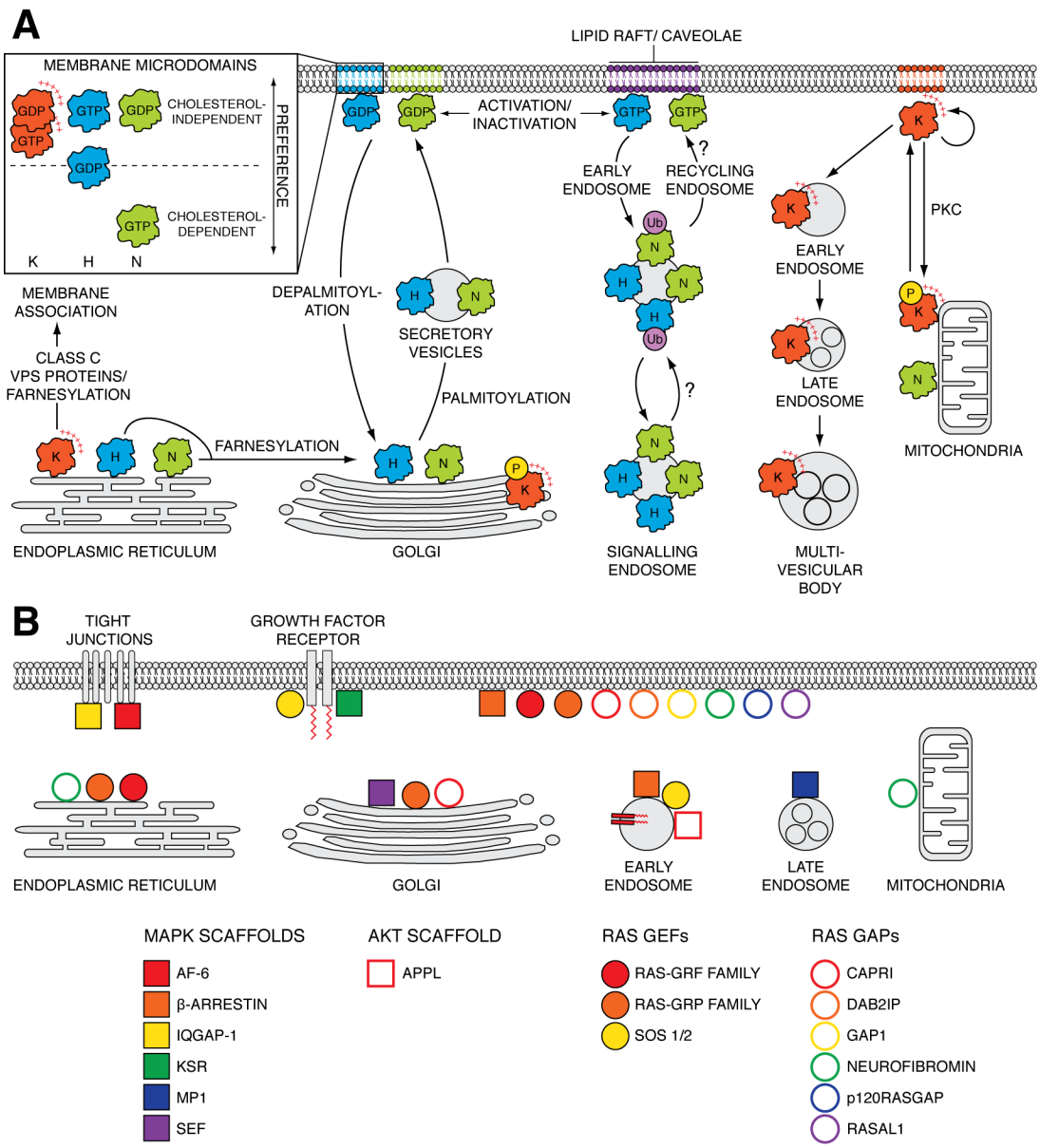


Fig. 1.6. Differential cellular localisations of Ras isoforms and compartmentalised signalling. (**A**) Localisation of Ras isoforms throughout the cell. H- and N-Ras share some degree of overlap in their localisation, whereas K-Ras exhibits the most distinct localisation. Inset shows Ras isoform nanocluster preference for cholesterol, which alters dependent on activation status. (**B**) Compartmentalisation of Ras GEFs, GAPs and scaffolding proteins. Most of the listed proteins exhibit regulated recruitment from the cytosol.

Furthermore, there is further lateral segregation dependent upon activation state, such that GTP and GDP-bound Ras are spatially distinct for each isoform (Prior et al., 2001; Prior et al., 2003). They also exhibit switch-like behaviour over a wide input range, which in effect coverts an analogue signal into a digital one (Harding and Hancock, 2008). Interestingly, if Ras nanocluster formation is inhibited, signalling output is reduced to 3% of maximum (Tian et al., 2007). Such distinct nanocluster formation is believed to further underlie Ras isoform-specific signalling.

1.1.4 RAS ABUNDANCE

Modern biology is nearing the end of identifying/ characterising individual proteins and entering a phase where it aims to understand how proteins work together to form molecular networks and circuits (Bordbar et al., 2014; Janes and Lauffenburger, 2013). Since many signalling pathways contain large numbers of components and our minds can only make simple connections between them, the behaviour of complex networks is often difficult to understand. As a result, mathematical modelling of pathways is becoming a popular means to interpret and interrogate large signalling maps, as it can combine many aspects of signalling networks and provide platforms to generate and test new hypotheses (Klipp and Liebermeister, 2006). Although pathway models can be created with unknown variables or incomplete data, it is important to provide constraints to make the model more accurate. One key constraint is that of component concentration (Aldridge et al., 2006) and despite being studied for three decades, it is still unclear how many Ras molecules are present in a cell. Therefore, models of the MAPK and PI3K pathways currently lack a critical parameter. Here the current knowledge of cellular Ras abundance is discussed.

The first data on Ras abundance was provided by RNA analysis of human and mouse samples (Fiorucci and Hall, 1988; Leon et al., 1987). These studies showed that, although H-, K- and N-Ras isoforms are ubiquitously expressed, RNA levels for each isoform vary considerably between tissue types and stages of development, e.g. K-Ras RNA was highly expressed in lung tumour samples, but was poorly expressed in mouse foetal liver samples, relative to H- and N-Ras. However, as RNA levels do not accurately predict protein expression (Gygi et al., 1999b), these

studies can only provide a provisional view on Ras protein abundance. Immunoblotting of commonly used cell lines, that utilised the change in signal of a pan-Ras antibody following isoform knock-down, demonstrated that K- and N-Ras are relatively the most abundant isoforms, while H-Ras is the minor isoform (Omerovic et al., 2008). In contrast, the identification of rare codons in the KRAS gene suggests that it is poorly translated due to the stalling of ribosomes (Lampson et al., 2013). Transfection of identically tagged H- and K-Ras cDNA, showed that H-Ras was ~20-fold better expressed; plus converting the rare codons to common ones in the KRAS gene increased the abundance of K-Ras to levels comparable with H-Ras. These results indicate that K-Ras could be the least abundant isoform. Consequently, it is clear that the conclusions regarding Ras isoform abundance from the Lampson and Omerovic studies do not match. Moreover, neither study specifies an exact number of Ras proteins per cell.

Presently, a popular approach to measure cellular protein abundances is by using mass spectrometry (MS)-based proteomics. Proteomics offers a variety of quantification strategies that can provide sensitive and accurate measurements of protein expression levels (see section 3.6 for a detailed discussion). Over the past few years, several targeted strategies have been applied to the question of cellular Ras abundance; see Table 1.1 for a summary. Vogelstein and colleagues extracted Ras isoforms from cell lysates using an immunoprecipitation strategy, before elution of bound Ras proteins and tryptic digestion into peptides. To achieve quantification, AQUA peptides (isotope-labelled peptide standards that act as surrogates for endogenous peptides) were then spiked into the digested sample followed by selected reaction monitoring (SRM) analysis on a triple quadrupole mass spectrometer (Wang et al., 2011). They found that on average there is over 1.3 million Ras molecules per cell and that K-Ras is the major isoform, while H- and N-Ras are expressed at similar levels. Although AQUA provides precise measurement, it does make assumptions regarding complete protein digestion and recovery that will impact upon its accuracy (Gallien et al., 2011; Lowenthal et al., 2014). Plus, in the Vogelstein study, a proportion of Ras protein remained attached to the beads after elution and that only ~25% of Ras from cell lysates was used for MS analysis, necessitating the use of a correction factor.

Study	Sample prep. and quantification technique	Cell line(s)/ sample	ng Ras per 100 μg sample	
			wt	Mutant
Halvey et al., 2012	SDS-PAGE, AQUA	CaCo2	5.12 ± 0.17	N/A
		HCT 116	2.35 ± 0.36	1.28 ± 0.19
		LoVo	1.49 ± 0.02	N/A
		LS174	N/A	1.28 ± 0.02
		SW480	N/A	6.82 ± 0.04
Wang et al., 2011	Ras immuno-precipitation, AQUA	SW480	17.05 ± 2.98	95.93 ± 22.38
		Pa08C	27.71 ± 5.22	43.70 ± 6.08
		Pa16C	19.19 ± 1.49	9.38 ± 0.22
		Pa02C	26.65 ± 1.49	0
		Spleen 10 (N)	3.73 ± 0.43	0
		Colorectal 102 (N)	4.37 ± 0.23	0
		CRC 2640 (T)	0.93 ± 0.33	0.64 ± 0.20
		CRC 2966 (T)	1.39 ± 0.18	0.37 ± 0.04
		Cyst 3950 (T)	0.18 ± 0.19	0.19
		Cyst 10592 (T)	0.16	0.19
Ruppen-Canas et al., 2012	Ras immuno-precipitation, AQUA	PANC198 (X)	0.0079 ± 0.0014	0.0075 ± 0.0002
		PANC354 (X)	0.015 ± 0.0012	0
		PT3092 (T)	0.0037 ± 0.0004	0.0013 ± 0.0008
		PT5778 (T)	0.0062 ± 0.00005	0.00093 ± 0.0002
		CRC010 (X)	0.0026 ± 0.0002	0.00113 ± 0.00002
		CRC020 (X)	0.0027 ± 0.00005	0
		CT5669 (T)	0.0029 ± 0.00015	0.0016 ± 0.00007
		CT5381 (T)	0.0022 ± 0.0013	0

Table 1.1. Summary of Ras abundance measurements from previous targeted proteomics studies. Total wild-type and mutant (G12V, G12D or G13D) abundances detected in various cell lines, normal tissues (N), tumour samples (T) and xenograft samples (X). Cyst, pancreatic cyst; PANC and PT, pancreatic tumour; CRC and CT, colorectal tumour. Data from each study were converted into ng per 100 μg sample. Halvey et al. performed 3 biological replicates for the HCT 116 cell line but 3 technical replicates for the remaining cell lines. Data presented as mean ± SD. n = 3 for all samples where SD given.

Ashman and co-workers employed the same approach to measure Ras abundance in tumour and xenograft samples (Ruppen-Canas et al., 2012). In one pancreatic tumour sample, wild-type and mutant G12D Ras were expressed at 37.2 ± 3.6 and 13.4 ± 8.2 pg/ mg tissue, respectively. However, when these values and those from the Vogelstein study are converted into the same units, it is clear that there is a significant disagreement despite both studies employing the same technique (Table 1.1). Liebler and co-workers employed SDS-PAGE to enrich for Ras proteins, then cut out the Ras-containing gel band and performed an in-gel digestion. Following

peptide extraction, AQUA peptides were added to enable quantification in SRM analyses (Halvey et al., 2012). They report that wild-type Ras is expressed at 0.7 ± 0.01 fmol/ μg in LoVo cells, 2.4 ± 0.08 fmol/ μg in Caco2, while mutant Ras has concentrations of 0.6 ± 0.01 fmol/ μg in LS174T cells (G12D/G13D) and 3.2 ± 0.02 fmol/ μg in SW480 cells (G12V). When these values are converted into ng per 100 μg lysate, they are around 10-fold less than the results of the Vogelstein study. However, it is unclear how efficient the in-gel digest is and how many Ras peptides remain trapped in the gel. Other proteomic studies, that adopt an entirely different approach as those mentioned above by employing a global, label-free quantification strategy, can provide a number of Ras molecules per cell. Selbach and colleagues employed intensity-based absolute quantification (iBAQ) to measure the abundance of over 6,000 proteins in mouse NIH-3T3 cells (Schwanhausser et al., 2011). The iBAQ algorithm normalises the summed intensities of observed peptides by the number of theoretically observable peptides from a protein to calculate protein copy numbers per cell (see section 3.6 for a more detailed explanation). In this study, Selbach reports that per NIH-3T3 cell the average protein copy number for H-Ras is 22,089, for K-Ras 41,089 and N-Ras 378,351. Utilising the same technique, the Mann laboratory quantified over 10,000 different proteins from the HeLa cervical cancer cell line and calculated that 173,933 K-Ras molecules are present per cell (Nagaraj et al., 2011). However, the iBAQ technique can only provide approximate protein copy numbers per cell, since the utilised algorithm predicts the number of theoretically observable peptides, which could be inaccurate for many proteins. Nevertheless, these numbers are in stark contrast to those reported in the Vogelstein study, where it is reported that per SW480 cell there is 1.5 and 8.6 million molecules of wild-type and mutant Ras, respectively (Wang et al., 2011). In summary, the techniques employed to date to quantify cellular Ras abundance have several flaws, including the assumption that all endogenous protein is fully digested (as is the case in AQUA, employed by Halvey et al., Wang et al. and Ruppen-Canas et al.), that all peptides are recovered from the following in-gel digestion (as in Halvey et al.) and inability to compensate for differences in MS responses between detected peptides (as in Schwanhausser et al. and Nagaraj et al.). Many of these challenges are overcome using the PSAQ technique, which ensures standard and endogenous proteins are equally digested, the identical peptides are used for quantification and that standard and endogenous peptides undergo the same losses from the beginning of the experiment.

1.2 RAS IN CANCER

Cancer is characterised by the uncontrolled division of abnormal cells in the body. Following the multistep process of carcinogenesis, cancerous cells sustain proliferative signalling, evade growth suppressors, exhibit replicative immortality, induce angiogenesis, resist cell death and may also metastasise and invade other tissues (Hanahan and Weinberg, 2000, 2011). Ras proto-oncogenes are frequently mutated in human cancer, with an average mutation incidence of 16% (Prior et al., 2012). However, analysis of isoform mutations reveals a clear bias in certain cancer types. Furthermore, each isoform mutation also has a distinct codon mutation signature and clinical data has demonstrated that specific codon mutations result in differential sensitivity to chemotherapeutic agents (De Roock et al., 2010; Mao et al., 2013), suggesting that different codon mutations are unlikely to be equal in their signalling output. Discussed here is the spectrum and bias of Ras isoform and codon mutations observed in human cancer, the biochemical and structural properties of various Ras mutants and the biological effects and clinical relevance of Ras mutations.

1.2.1 INCIDENCE AND SPECTRUM OF RAS MUTATIONS IN CANCER

Out of the 3 Ras proto-oncogenes, K-Ras is the by far the most commonly mutated in human cancer. According to the Catalog of Somatic Mutations in Cancer (COSMIC), K-Ras was mutated in 22% of all tumours analysed, compared to 8% for N-Ras and 3% for H-Ras (Cox and Der, 2010; Prior et al., 2012). At a glance these numbers imply that Ras is mutated in ~30% of cancers, but, when all cancers with more than 20 tumours analysed were given equal weighting, the actual mutation rate drops to 16%. In most cancer types, there is a clear preference for mutation of a single isoform, e.g. pancreatic cancer has K-Ras mutated in 61% of cases, N-Ras in 2% and H-Ras 0%; colorectal cancer (CRC), K-Ras 33%, N-Ras 3% and H-Ras <1%; skin cancer, N-Ras 18%, H-Ras 6% and K-Ras 3% (Prior et al., 2012). One exception is thyroid cancer, where the mutation rate is more equal: N-Ras 8%, H- and K-Ras both 3%. These mutation patterns are highly unlikely to occur by chance and suggest that certain Ras mutations happen for a reason in particular cancer types. In support of this notion, experiments using genetically engineered mice showed that oncogenic K-Ras, but not N-Ras, promoted colonic tumorigenesis

(Haigis et al., 2008). Such results demonstrate that there are clear phenotypic differences between Ras isoform mutants.

Similar to the observed isoform bias in specific cancers, each Ras isoform has a preference for certain codon mutations. While Ras can be activated by single-base substitutions at codons 12, 13 or 61, 80% of K-Ras mutations occur at codon 12 and almost 60% of N-Ras mutations at codon 61 (Prior et al., 2012). H-Ras mutations are split 50% and 40% between codons 12 and 61, respectively. Such differences in codon specificity are unexpected, as these mutations fall in regions where Ras isoforms are identical and therefore assumed to have the same effects on protein activity; plus, K-Ras and N-Ras genes share identical sequences encoding G12 and Q61. Although codons 12, 13 and 61 can each be converted by base substitutions to encode 6 different amino acids, over 60% of Ras isoform mutations are accounted for by 3 of the possible 18 mutations (Cox and Der, 2010; Prior et al., 2012). K-Ras mutations are dominated by G-to-A transitions at codons 12 or 13, which give rise to G12D and G13D, respectively, and G-to-T transversions at codon 12 to produce G12V. N-Ras exhibits a similar pattern of codon 12 and 13 mutations, except at a much reduced rate, but also suffers from mutations in codon 61 giving rise, for example, to Q61K, Q61L and Q61R. Interestingly, K-Ras and N-Ras G13D mutations are only detected in significant numbers in colorectal and lymphoid cancers, respectively. In contrast, H-Ras favours G12V in cancers with codon 12 mutations and shares a similar spectrum of codon 61 mutations as N-Ras. Together, these observations show that there is a link between Ras isoform and codon mutational status with specific cancer types.

1.2.2 ONCOGENIC RAS MUTANTS

Ras GTPases with mutations in positions 12, 13 or 61 are constitutively activated and signal uncontrollably. Each of these mutations interfere with the catalysis of GTP hydrolysis, causing Ras to remain locked in its active, GTP-bound form (Prive et al., 1992). Position 12/ 13 mutants are also resistant to GAP activity (Trahey and McCormick, 1987), as substitution of the glycine residues at either of these positions with any other amino acid, even alanine, will sterically clash with a GAP's arginine finger and prevent it from entering the Ras nucleotide-binding pocket to accelerate GTPase activity (Fig. 1.7) (Scheffzek et al., 1997). Other than proline, any amino

acid will activate Ras if it is substituted at position 12 (Franken et al., 1993; Seeburg et al., 1984). Position 61 mutants lose the catalytic glutamine residue essential for GTP hydrolysis. In wild-type Ras, Q61 is orientated following GAP binding to optimally align a water molecule and enable the GTP hydrolysis reaction (Fig. 1.2) (Milburn et al., 1990). Any amino acid, bar proline, glutamic acid and glutamate, substituted at position 61 confers transforming activity (Der et al., 1986; Polakis and McCormick, 1993). In the following section the structural differences between various mutant Ras proteins are discussed, along with their biochemical and biological properties.

1.2.2A STRUCTURAL AND BIOCHEMICAL PROPERTIES OF RAS MUTANTS

Wild-type and mutant Ras proteins as a whole are extremely similar, but have noticeable structural differences located around the nucleotide-binding pocket (Fig. 1.7). G12D mutants exhibit a network of aberrant hydrogen bonding between the mutant aspartate residue, Y32 and G60, which causes P34 to become fully exposed at the molecule's surface (Al-Mulla et al., 1999; Franken et al., 1993). In G12V and Q61L mutants, Y32 is orientated slightly differently (Buhrman et al., 2011; Buhrman et al., 2007). It has also been suggested that two pocket regions, located near the Ras binding site, are structurally different between wild-type and G12V Ras (Muraoka et al., 2012). Ras mutants also display differences in their conformational states. When GTP bound, Ras assumes two interconverting conformations, 'inactive' state 1 and 'active' state 2, which are characterised by different positioning of Y32 relative to the nucleotide (Geyer et al., 1996; Shima et al., 2010). State 2 closely resembles the Ras effector-bound state and has high affinity for effectors, as well as increased GTPase activity, while state 1 has reduced effector affinity and GTPase activity (Muraoka et al., 2012). Wild-type Ras can alternate between these states every few milliseconds, with state 1 population occupying 36 ± 2% for wild-type H-Ras-GppNHp (Shima et al., 2010; Spoerner et al., 2004). However, mutations shift the equilibrium between the two states, as state 1 occupies 53% and 58% for H-RasG12V-GppNHp and H-RasQ61L-GppNHp, respectively (Spoerner et al., 2004; Ye et al., 2005). The G12D mutation has also been shown to stabilise the state 1 conformation (Geyer et al., 1996). However, it is uncertain how the equilibrium is shifted in many of the other Ras mutants and the biological relevance of such shifts is unclear.

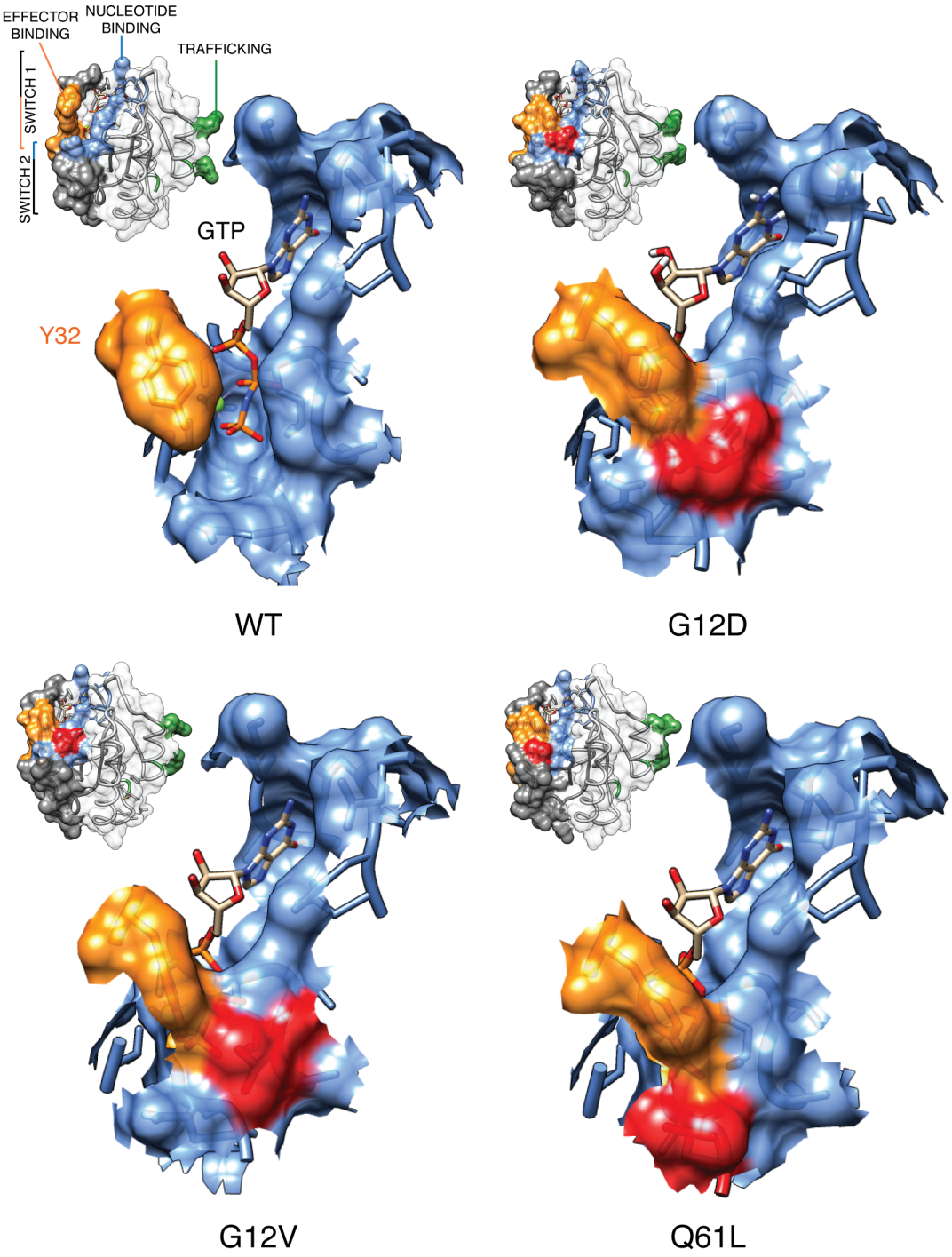


Fig. 1.7. The structure of oncogenic Ras. Whole molecules and nucleotide-binding pockets of wild-type and oncogenic H-Ras mutants: G12D, G12V and Q61L. All Ras structures are in their active conformation, bound to GppNHp. Mutation highlighted in red. WT H-Ras, PDB code 1ctq (Scheidig et al., 1999); G12D, PDB code 1agp (Franken et al., 1993); G12V, PDB code 3oiw (Buhrman et al., 2011); Q61L, PDB code 2rgd (Buhrman et al., 2007).

The above-mentioned structural differences may account for the discrete biochemical differences observed between mutant proteins, because specific mutations at position 12, 13 or 61 are likely to disturb residues within the binding pocket in their own unique way (Scheffzek et al., 1997). Oncogenic Ras mutants exhibit various dissociation rates for guanine nucleotides, e.g. GTP dissociation from G12D mutants is 8-fold faster than wild-type or G12V Ras (John et al., 1988; John et al., 1989). Ras mutants also display differentially impaired intrinsic GTPase activities compared to wild-type, e.g. Trahey et al. reported that H-Ras G12V and G12D retain 12% and 43% of wild-type GTPase activity, respectively (Al-Mulla et al., 1999; Colby et al., 1986; Der et al., 1986; Franken et al., 1993; Gibbs et al., 1984; John et al., 1988; Krengel et al., 1990; Trahey et al., 1987). However, the literature on mutant GTPase activity doesn't entirely agree, possibly because different assay techniques were used. Recently, it has been reported that intrinsic Ras GTPase activity can be increased following Raf-1 binding via allosteric modulation (Buhrman et al., 2010). Ras mutants may not to be sensitive to this emerging mode of regulation, as their impaired GTPase activity may be non-enhanceable and mechanistically their mutated binding sites may not be unable to support it. Consequently, Ras mutations may provide a means to circumvent this effector-mediated regulation.

1.2.2B DIFFERENT TRANSFORMING ACTIVITIES AND SIGNALLING OF RAS MUTANTS

Focus formation assays have demonstrated that different H-Ras mutations exhibit a range of transforming abilities, e.g. G12V mutants exhibited strong transforming effects, while G12D, G12C, G12A and G13N transforming activities were slightly weaker (Der et al., 1986; Fasano et al., 1984; Seeburg et al., 1984; Smith et al., 2010). These differences are very likely attributable to the subtle, but unique, biochemical and structural variations of each mutant Ras protein, as detailed above. Interestingly, G12P mutations, although insensitive to GAPs, could not transform cells, since it retains intrinsic GTPase activity (Schweins et al., 1995). Furthermore, K-Ras G12V transfectants have been shown to induce tumours in nude mice in shorter time and latency compared to G12D and G12C (Cespedes et al., 2006). Together, these studies show that different Ras mutations, even if they affect the same codon, cannot be considered the same, since they must signal differently in order to exert different biological effects.

The position of the Ras point mutation also seems to be associated with the degree of tumour cell aggressiveness. Codon 12 Ras mutations exhibit increased transforming capacity compared to those at codon 13 (Bos et al., 1985; Fasano et al., 1984; Sloan et al., 1990), offering a possible explanation why these codon mutations are more frequently observed in human cancers (Capella et al., 1991; Prior et al., 2012). Stable expression of K-RasG12D and K-RasG13D in NIH 3T3 cells showed several functional differences, e.g. G12D mutant cell line formed colonies of increased density, spontaneously underwent anchorage-independent growth and had reduced ability to trigger apoptosis compared to G13D mutant cells (Guerrero et al., 2000). Guerrero et al. also show that G12D and G13D had similar abilities to activate ERK, but the former had greater capacity to activate AKT, suggesting that G12D signals more strongly through the PI3K pathway. This observation has, in part, been observed in nude mice xenografts, where K-RasG12D demonstrated increased activation of AKT but very little activation of ERK (Cespedes et al., 2006). Structural analysis of the K-RasG12D protein predicts that it cannot interact with ERK, possibly explaining the reduced ERK activation (Al-Mulla et al., 1999). Injection of nude mice with NIH 3T3 transfectants showed that K-RasG12D induced fibrosarcoma-like tumours, while K-RasG13D created tumours that resemble fibrous histocytomas (Guerrero et al., 2002). This study also showed that tumours derived from G13D transfectants grew at a slower rate and exhibited a higher level of apoptosis compared to the G12D transfectants. Plus, G13D tumour-derived cells expressed more K-Ras and exhibited lower levels of AKT activation compared to those with G12D. Analysis of the metabolic flux of transfected NIH3T3 cells using [1,2-$^{13}C_2$] glucose showed that K-Ras-G12D and -G13D mutations produce distinctive metabolic phenotypes (Vizan et al., 2005). Cells harbouring K-RasG12D utilised more glucose for anaerobic glycolysis compared to G13D, which could explain how G12D cells have increased resistance to apoptosis, while G13D demonstrated increased use of anabolic pathways.

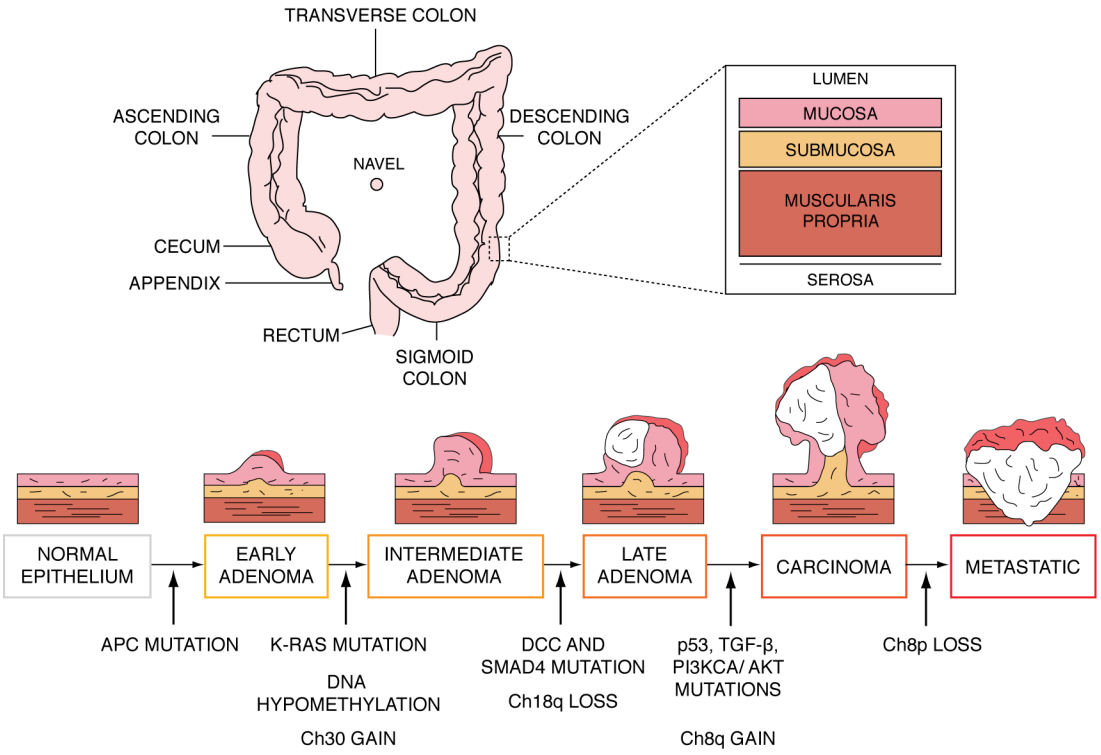


Fig. 1.8. Colorectal carcinogenesis. (Top) Anatomy of the colon and tissue layers that compose the colon wall. (Bottom) Vogelgram detailing the genetic alterations and mutations observed at different stages of colon cancer development. Based on (Fearon and Vogelstein, 1990; Saif and Chu, 2010).

1.2.3 CLINICAL RELEVANCE OF RAS CODON MUTATIONS

CRC is the third most common cancer worldwide with an overall 5-year survival rate of 11% for patients exhibiting metastatic disease (Siegel et al., 2012). In CRC, K-Ras is most commonly activated at codon 12 (33% of cases), with aspartate (40% of codon 12) and valine (37%) amino acid substitutions the most frequent, followed by cysteine (11%) and serine (7.7%); codon 13 mutations are less prevalent (9.5% of cases) but aspartate substitutions dominate this position (95%) (Vaughn et al., 2011). This mutation usually occurs at an early stage of colon carcinogenesis (Fearon, 2011; Fearon and Vogelstein, 1990; Vogelstein et al., 1988) (Fig. 1.8). The prognostic value of K-Ras mutations in CRC is uncertain, since many studies make conflicting conclusions (Anwar et al., 2004; Cerottini et al., 1998; Lievre et al., 2006; Locker et al., 2006; McLeod and Church, 2003; Ren et al., 2012). However, comparing survival statistics of patients who harbour specific Ras mutations reveal differences in clinical outcome. The Kirsten ras in-colorectal-cancer collaborative group (RASCAL) studies, which monitored over 3,000 patients from several countries, show that K-RasG12V mutations are associated with poorer prognosis

than other mutations, i.e. shorter disease-free survival and survival rate (Andreyev et al., 2001; Andreyev et al., 1998). However, other studies found that K-RasG13D mutations were associated with a worse clinical outcome. Bazan et al. monitored 160 Italian CRC patients and found that those with codon 13 mutations had worse disease-free and overall survival compared to individuals with codon 12 mutations (Bazan et al., 2002). Similarly, Heinemann and co-workers, who studied 273 German patients, found that those with a K-RasG13D mutation presented a more aggressive disease associated with local and distant metastasis at first diagnosis (Modest et al., 2011). Samowitz et al., who reviewed 1,413 American patients, found that K-RasG13D mutations were associated with a 40% increase in risk of short-term mortality and Cerottini et al. showed that these mutations are linked with reduced disease free survival following surgery (Cerottini et al., 1998; Samowitz et al., 2000). Yet, Ogino and colleagues mined two pathological epidemiology databases of 1,261 rectal and colon tumours and found that G12D mutations, not G13D, were associated with inferior survival in BRAF wild-type CRC patients (Imamura et al., 2012). Overall, it is clear that some disagreement exists between studies, but the majority indicate that codon 13 mutations predict the worst patient outcome and out of the codon 12 mutations, valine substitutions may confer more aggressive tumour behaviour.

Cetuximab (Erbitux®; ImClone systems/ Eli Lilly/ Bristol Myers Squibb), an immunoglobulin G1 monoclonal antibody against EGFR, has proven effective as a single agent or in combination with chemotherapy for the treatment of patients with metastatic CRC (De Roock et al., 2008; Jonker et al., 2007). This antibody binds the to the extracellular domain of EGFR to competitively inhibit EGF or other growth factors from binding, which prevents receptor phosphorylation and subsequent activation of downstream signalling events (Baselga, 2001; Graham et al., 2004). However, the European Medicines Agency and US Food and Drug Administration advise that patients harbouring codon 12 or 13 K-Ras mutations should be excluded from cetuximab treatment, since these mutations have been reported as resistance markers to anti-EGFR therapy (Karapetis et al., 2008; Lievre et al., 2008; Lievre et al., 2006). Yet, previous studies showed that a few patients with K-Ras mutant tumours did occasionally respond to anti-EGFR treatment (Frattini et al., 2007; Moroni et al., 2005). The majority of these patients suffered from K-RasG13D mutations, which suggests that codon 12 and 13 K-Ras mutations may have

different effects on cetuximab sensitivity. A retrospective analysis of 579 CRC patients treated with cetuximab, between 2001 and 2008, revealed patients harbouring a K-RasG13D mutation were associated with a longer overall survival (OS) and progression free survival (PFS), 7.6 months and 4.0 months, respectively, when treated with cetuximab monotherapy or any cetuximab-based therapy, compared to patients with other K-Ras mutations, OS 5.7 and PFS 1.9 months (De Roock et al., 2010). This study also demonstrated that isogenic colorectal cell lines containing wild-type or G13D K-Ras were sensitive to cetuximab treatment, while cells with K-RasG12V were resistant (Di Nicolantonio et al., 2008). Similarly, Messner et al. used a panel of cell lines harbouring various K-Ras mutations to show that those harbouring G13D mutations were more susceptible to cetuximab treatment compared to those with mutations at codons 12 or 61 (Messner et al., 2013). Tejpar et al. analysed 533 CRC patients, of which 83 had G13D mutations and benefitted from first-line treatment with cetuximab (Tejpar et al., 2012). Cetuximab plus chemotherapy versus chemotherapy alone improved PFS (7.4 versus 6 months) for G13D patients, but had little benefit for OS (15.4 versus 14.7 months). Tejpar and colleagues also noticed that G13D patients receiving chemotherapy alone had a worse outcome compared to patients with other K-Ras mutations. Bando et al. retrospectively analysed 109 Japanese CRC patients, who were treated with irinotecan and cetuximab and found that the seven G13D patients in the analysis did not have significantly improved OS or PFS compared to other patients, despite showing a higher disease control rate (Bando et al., 2012). Gajate et al. reviewed 110 Spanish CRC patients and showed that G13D patients had no statistically different OS or PFS compared to other mutations (Gajate et al., 2012). However, neither the Gajate or Bando studies had G13D control groups, as all patients received cetuximab treatment in these retrospective analyses.

Patients who initially respond to cetuximab treatment often develop secondary resistance following several months of treatment (Karapetis et al., 2008). Exposing cells to increasing concentrations of cetuximab *in vitro* causes them to develop acquired resistance, either through amplification of K-Ras alleles and increased protein expression or the appearance of K-Ras mutations (Diaz et al., 2012; Misale et al., 2012). Chronic treatment of cells with MET or MEK inhibitors also show that acquired resistance can occur through amplification of K-Ras (Cepero et al., 2010; Little et al., 2011). However, screening of 1,039 CRC samples revealed that K-Ras

amplification is a rare event occurring in 0.67% of samples; plus K-Ras amplification and K-Ras mutations seem to be mutually exclusive (Valtorta et al., 2013). Patients are likely to develop resistance because all tumours, even those characterised as K-Ras wild type, are likely to have subpopulations of K-Ras mutant cells that DNA sequencing simply isn't sensitive enough to detect (Parsons and Myers, 2013). Consequently, when patients are treated with cetuximab, the outgrowth of these mutant subpopulations occur leading to anti-EGFR resistance.

1.3 MASS SPECTROMETRY-BASED PROTEOMICS

1.3.1 OVERVIEW OF MASS SPECTROMETRY

Mass spectrometry (MS) is an analytical technique capable of measuring atoms and molecules, in their ionised form, to determine their molecular weight with great accuracy. It has a broad range of applications in such fields as food and water quality control, sports drug testing, forensic sciences, atomic physics and many others (de Hoffmann and Stroobant, 2013). Owing to technological advances over the past 15 years, in particular with ionisation, MS has most recently been applied to biological questions, such as the large-scale study of proteins, known as proteomics (Wilkins et al., 1997). Proteomics is one of many '–omics' (e.g. metabolomics, genomics, transcriptomics) that collectively aim to characterise how biological molecules work in concert to understand the function and dysfunction of whole organisms. MS-based proteomics was the central technology employed in this thesis to address the questions of cellular Ras abundance and mutant Ras codon-specific signalling. A mass spectrometer, the analytical instrument used in MS, manipulates and measures ionised species in the gas phase at low pressures by the application of electric and/or magnetic fields. A typical mass spectrometer consists of three basic components: an ion source, a mass analyser and a detector (Watson and Sparkman, 2007). Each of these components is further discussed with a particular focus on mass analysers. Also detailed are the two mass spectrometers employed in this thesis: the 4000 QTRAP and LTQ Orbitrap XL.

1.3.2 TYPICAL MASS SPECTROMETRY-BASED PROTEOMICS WORKFLOW

Accurately measuring the mass (or more precisely the mass-to-charge ratio, *m/z*) of biological molecules is a straightforward task for a modern mass spectrometer. In high-throughput proteomics, peptides rather than proteins are analysed and these are used to infer information regarding the proteins (a strategy known as surrogacy). This well-established 'bottom up' approach is favoured for numerous reasons, e.g. proteins are challenging to handle and might not be soluble in the same buffer, plus the sensitivity of the mass spectrometer is lower for proteins versus peptides (Steen and Mann, 2004). Crucially, sequence information is required if the experimental aim is to identify proteins. The mass spectrometer is most efficient at providing such

information using peptides ~20 residues long, not whole proteins. Yet, it must be noted that analysing intact proteins with MS ('top down' approach) is a growing field with potentially exciting applications, e.g. improved analysis of PTMs (Doerr, 2008); however, it is beyond the scope of this thesis. In a typical GeLC-MS 'bottom up' MS-based workflow, proteins are isolated (released) from a biological source, e.g. cell culture or animal tissue (Fig. 1.9, step 1), fractionated using sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) (step 2) before in-gel proteolytic digestion (step 3) (Jeno et al., 1995; Shevchenko et al., 1996; Wilm et al., 1996), typically with trypsin (Olsen et al., 2004). Despite protein digestion resulting in a considerable increase in sample complexity, most proteins yield at least some soluble and ionisable peptides, essential requirements for the subsequent MS analysis. A digested sample may contain thousands of peptides, spanning several orders of magnitude in abundance, which are impossible to effectively analyse simultaneously. Peptides are therefore commonly injected onto a high-performance liquid chromatography (HPLC) column directly coupled online to the mass spectrometer (LC-MS) (step 4) (Steen and Mann, 2004). C_{18} reversed-phase (RP) nanoflow HPLC provides high resolving power in the separation of peptide mixtures for MS (Makarov and Scigelova, 2010; Shen et al., 2002; Wolters et al., 2001). RP-HPLC comprises of two phases: a stationary and a mobile phase, between which the analytes (peptide) are partitioned. The stationary phase is typically formed of C_{18}-bonded silica particles resulting in a hydrophobic surface. The mobile phase is composed of a polar (e.g. water) and organic (e.g. acetonitrile) solvent mixture at low pH (pH < 3). Formic acid in the mobile phase causes low pH, which ensures peptides are positively charged and therefore improves peptide retention and aids in electrospray ionisation (discussed below), plus the low pH minimises ionised silica groups that may reduce resolving power. At the start of a typical gradient, the mobile phase is highly polar (> 95% aqueous) resulting in the binding of peptides through hydrophobic interactions with the C18 phase. As the proportion of organic solvent in the mobile phase increases, peptides elute in order of their hydrophobicity, delivering much simpler peptide mixtures for MS analysis (Biosciences, 1999; Shi et al., 2004).

Peptides elute from the HPLC column at the ion source region of the mass spectrometer. Because MS measurements require gaseous ions, eluting peptides are electrostatically dispersed from a glass or metal capillary at the end of the

column (step 5). Known as electrospray ionisation (ESI), this technique was originally developed by Malcolm Dole (Dole et al., 1968) before John Fenn applied it to proteins (Fenn, 1993; Whitehouse et al., 1985). Fenn was awarded the 2002 Nobel Prize in Chemistry, along with Koichi Tanaka who developed matrix-assisted laser desorption/ionisation (MALDI) (Liu et al., 1981; Tanaka et al., 1988). Both ESI and MALDI are 'soft' ionisation techniques as they overcome the propensity of macromolecules to dissociate when ionised. ESI is generated by applying a voltage (3-6 kV), under atmospheric pressure, to a liquid passing through a capillary at low flow rate. The electric field between the capillary tip and counter electrode, separated by 0.3-2cm, induces charge accumulation at the liquid surface. Conflicting forces between surface tension, which forms a droplet, and electrostatic attraction towards the counter electrode, causes the liquid to become cone shaped, known as the Taylor cone (Taylor, 1964). Positively charged droplets break from the Taylor cone to form a fine spray when a voltage threshold, the Rayleigh limit, is reached (Wilm and Mann, 1994). A curtain of heated inert gas (usually nitrogen) and/or a coaxial gas flow aids droplet desolvation. As droplets shrink their charge density increases because the liquid evaporates as a neutral species. These smaller, highly charged droplets undergo coulombic fission to form nanodroplets that lead to gaseous ions (de Hoffmann and Stroobant, 2013). Several ESI variations have been developed in recent years. One of note is nanoelectrospray ionisation (nESI) (Wilm and Mann, 1996), where nL/min HPLC flow rates are used instead of the μL/min of ESI. nESI generates smaller droplets that improves desolvation, enhances ionisation efficiency and therefore increases sensitivity. Typical flow rates of 200-400 nL/min are employed in modern nanospray systems for proteomics. Following ionisation, gaseous peptide ions enter the mass spectrometer and are guided towards the mass analyser by electrostatic lenses (step 6). Peptide analysis in the mass spectrometer is detailed in the following section, but briefly the mass analyser first measures the *m/z* of intact peptide ions (precursor ions), before a peptide of interest is isolated and fragmented; a strategy termed tandem MS (MS/MS) (de Hoffmann and Stroobant, 2013; Watson and Sparkman, 2007). Utilising the precursor and fragment ion *m/z* values, the amino acid sequence can be deduced. Working alongside the mass analyser is the detector, which has an essential role in detecting ions and measuring their intensities. Most often, peptide quantification and peptide-to-spectrum mapping are performed computationally (step 7).

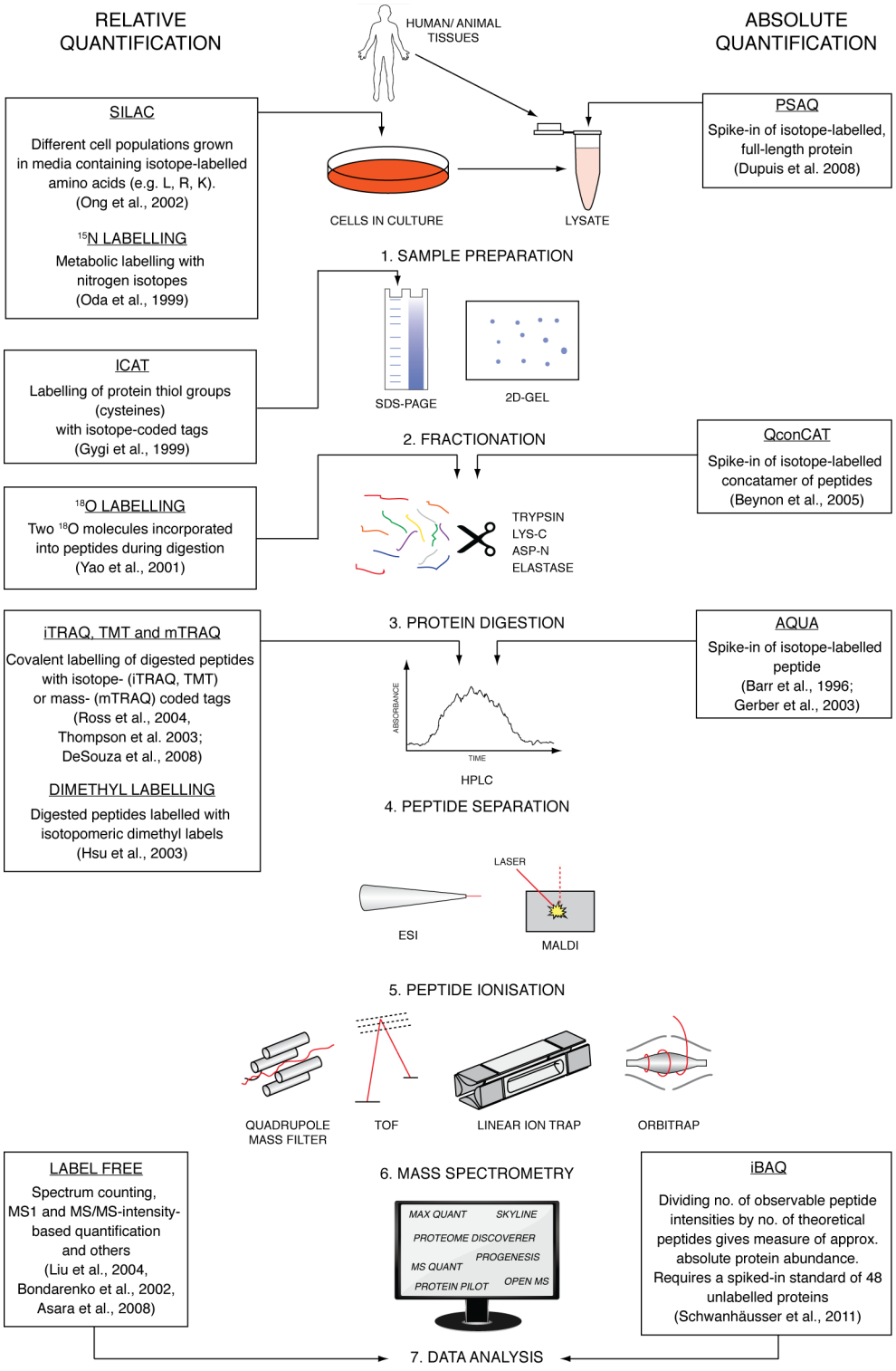


Fig. 1.9. Workflow of a typical MS-based proteomic experiment and examples of quantification strategies. Following sample preparation (1), fractionation is usually performed to reduce complexity (2) before digestion with trypsin or other protease (3). The subsequent peptide mixture is separated using reversed-phase high performance liquid chromatography (RP-HPLC) (4) then ionised using electrospray (ESI) (5). Alternatively, ionisation is achieved with matrix-assisted laser desorption/ionisation (MALDI). MS analysis, e.g. tandem MS, is performed in the mass spectrometer (6). Finally, peptide quantification and peptide-to-spectrum mapping are performed computationally (7). Arrows indicate the point reagents for a quantification method are added to the workflow.

1.3.3 MASS ANALYSERS

Once gaseous ions enter the mass analyser they are separated according to their *m/z*. Several types of mass analyser have been developed that utilise different principles to separate ions, but they all employ static or dynamic electric fields, either in combination or alone, to achieve ion separation. The ways these fields are employed represent some of the most fundamental differences between analyser types. Examples of mass analysers include the quadrupole mass filter (Paul et al., 1958), 3D ion trap (Steinwedel and Paul, 1960), linear ion trap (Hager, 2002; Schwartz et al., 2002), Time-of-Flight (TOF) (Wolff and Stephens, 1953), Fourier Transform Ion Cyclotron Resonance (FTICR) (Comisarow and Marshall, 1974) and Orbitrap (Makarov, 2000). What follows is a description of mass analysers employed in this thesis: the linear quadrupole, which forms the basis of the quadrupole mass filter (QMF) and linear ion trap (LIT), and the Orbitrap.

1.3.3A LINEAR QUADRUPOLES

Linear quadrupoles are possibly the most extensively employed components of mass spectrometers today; having multiple applications as ion guides, mass filters, collision cells and LITs (Douglas, 2009). Wolfgang Paul and co-workers described the principles of using quadrupole electric fields to manipulate ion trajectories in 1953 (Paul and Steinwedel, 1953), then developed the linear quadrupole for MS in 1958 (Paul et al., 1958). The term quadrupole is derived from the use of four opposing electrical poles to create an electric field. As such, linear quadrupoles are constructed of four circular or, preferably, hyperbolic-shaped rods aligned perfectly parallel (Fig. 1.10A) (de Hoffmann and Stroobant, 2013). The specific electrical potentials applied to the rods determine the function of the quadrupole. Discussed below is how these are employed as QMFs and LITs.

1.3.3B QUADRUPOLE MASS FILTERS

A QMF, also referred to as a transmission quadrupole, employs oscillating electric fields to selectively stabilise or destabilise ion trajectories according to their *m/z* value. Simply, these devices filter out a single *m/z* range from an unresolved ion beam (de Hoffmann and Stroobant, 2013). One pair of the QMF rods has an

oscillating radio frequency (RF) voltage superimposed on a positive DC voltage, while the other pair has a negative DC voltage and an RF voltage that is 180° out of phase with the first rod pair, which creates a quadrupole electric field in the centre of the rods (Fig. 1.10A). When the RF voltage reaches its peak opposite to the DC, it will overcome the DC and cause the rod to change polarity. The polarity is reversed when the RF voltage swings to its opposite peak. As typical RF voltages have constant frequencies of ~1.0 MHz, the rods oscillate their polarity hundreds of thousands of times every second (Douglas, 2009; Watson and Sparkman, 2007).

The basic principles of ion motion in a quadrupole field can be described with the Mathieu equation:

$$\frac{d^2u}{d\xi^2} + (a_u - 2q_u cos2\xi)u = 0$$

where u represents x, y and z axes, a_u and q_u are dimensionless trapping parameters, and ξ is another trapping parameter equal to $\Omega t/2$ (Ω is frequency and t time).

The Mathieu equation was devised by Émile Léonard Mathieu (1835-1890), to calculate stability regions of drum skins when struck (March, 1997), and utilised by Paul to encompass quadrupoles (Douglas, 2009; Paul et al., 1958):

$$\mathrm{U} = a_u \frac{m}{z}\frac{\omega^2 r_0^2}{8e} \qquad\qquad \mathrm{V} = q_u \frac{m}{z}\frac{\omega^2 r_0^2}{e}$$

where U is DC voltage; V, RF amplitude; *m/z*, mass-to-charge ratio; e, elementary charge (1.602×10^{-19} C); r_0, radius of the quadrupole; ω, $2\pi v$ (v is RF frequency in Hz). This 6-dimensional problem of ion transmission can be mathematically reduced to 2-dimensional plots represented as a_u, q_u space (Fig. 1.10B).

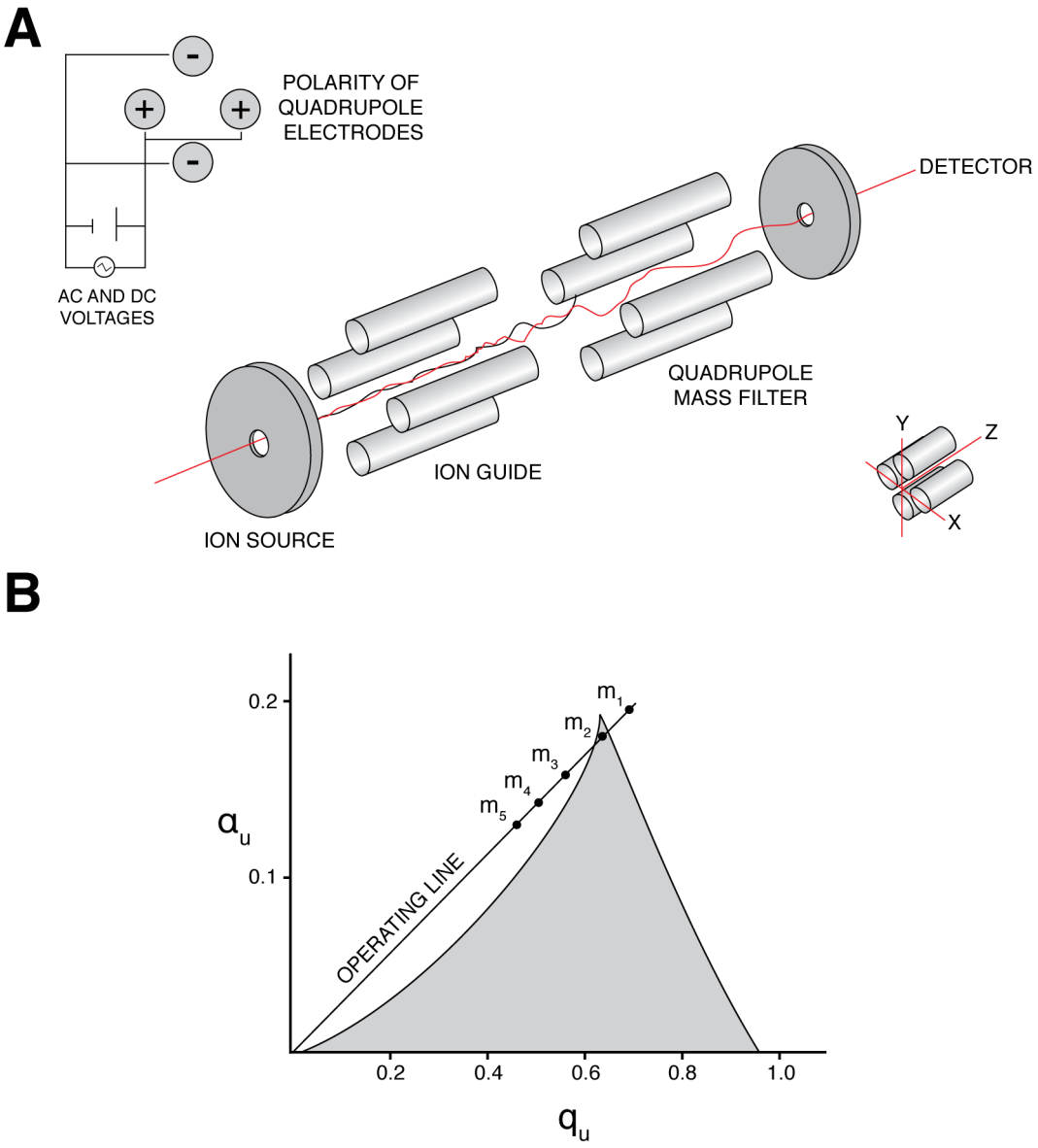


Fig. 1.10. Quadrupole mass filter (QMF) mass spectrometry. (A) Construction and operation of the QMF. Ions from the source pass through an RF only ion guide before reaching the QMF, which is constructed of four perfectly parallel rods with DC and RF potentials applied, that select for a single *m/z* value (resolution of 0.3 *m/z* units). Only ions with a stable trajectory (red) are transmitted to the detector, whereas unstable ions (black) strike a rod and are ejected from the quadrupole. (B) The first stability region. As RF and DC potentials are typically held at a fixed ratio, only values that lie along the operating line are available on the QMF. Grey area indicates stable region, occupied by theoretical ion m_2. RF and DC potentials must be increased or decreased for other theoretical ions to experience a stable trajectory and be transmitted. DC and RF amplitudes are related to a_u and q_u, respectively. Based on Douglas, 2009 and Watson and Sparkman, 2007.

When ions enter the QMF, they experience a total electrical quadrupolar field and must oscillate with a stable trajectory to be transmitted. An ion's trajectory is considered stable if the amplitude of its oscillations remains constant. Conversely, an ion with unstable trajectory has oscillations that increase exponentially with time, consequently the oscillation amplitude will eventually reach r_0 causing the ion to strike a rod, discharge and be ejected from the mass spectrometer (Fig. 1.10A) (Douglas, 2009). Transmission of ions with a single *m/z* value is achieved by scanning the amplitude of RF and DC voltages at a constant ratio (typically 6:1) to create an operating line; altering this ratio affects instrument resolution or sensitivity. At a certain RF/DC amplitude ratio, ions of a particular *m/z* will experience a stable trajectory and are transmitted, illustrated in Fig 1.10B. Overall, the QMF is a low-resolution instrument (0.3 Th), operates at constant separation across the *m/z* range, has a dynamic range of 10^5 and mass accuracy of 100 ppm (Douglas, 2009; Watson and Sparkman, 2007).

1.3.3C LINEAR ION TRAPS

Originally pioneered by Paul and Steinwedel, the quadrupole ion trap was originally designed as the 3D ion trap; a device (roughly the size of a tennis ball) constructed of two end cap electrodes and a ring electrode (Steinwedel and Paul, 1960). For his work in the development of the ion trap and QMF, Paul was awarded the 1989 Nobel Prize in Physics. However, following the introduction of the LIT in 2002 (Hager, 2002; Schwartz et al., 2002), with its 10-times improved ion trapping efficiency over the 3D trap (50% compared to 5%) and increased ion storage capacity, the 3D ion trap has effectively been replaced (Watson and Sparkman, 2007). Yet, the principles of ion motion in both instruments are essentially the same.

The LIT is based on the geometry of the QMF and confines ions radially by a 2D RF field and axially by DC stopping potentials applied to two end electrodes/ lenses (Fig. 1.11A). Like the QMF, the principles of ion motion in the LIT are explained by Mathieu equations, which can be mathematically reduced to 2D plots expressed as a_u, q_u space (Fig. 1.11B). Ions are only stored in the trap if their trajectories are stable in r- and z-directions simultaneously (r, radial distance to a rod; z, axial distance to an end electrode). When experiencing a stable trajectory in a 3D ion trap, ions follow a figure-of-eight shape or Lissajous curve (Fig. 1.11C) (Nappi et al.,

1997); this shape is elongated for a LIT. Ions can be ejected from quadrupole ion traps in a mass-selective manner. Ramping the RF voltage and/or applying an auxiliary RF voltage performs sequential mass-selective axial ejection (Hager, 2001; Londry and Hager, 2003). As the RF voltage is ramped, all ions experience an increase in axial velocity and when an ion's q_z reaches 0.908 (q_{eject}) it becomes unstable and is ejected. q_z is a dimensionless trapping parameter determining ion stability:

$$q_z = \frac{8\,eV}{m(r_0^2 + 2z_0^2)\omega^2}$$

where m is mass; z, charge; e, elementary charge (1.602×10^{-19} C); V, RF amplitude; r_0, radius of the quadrupole ; ω, $2\pi v$ (v is RF frequency in Hz). As ions with low *m/z* values have a higher initial q_z they are ejected first. A useful analogy for this is an exotic drink arranged of several coloured liqueurs; if the drink is poured, the lightest liqueurs of the drink will be poured out first, followed by the second lightest and so on. In resonance excitation ejection, an additional RF voltage is applied to the end electrodes causes specific ions to obtain increased axial velocity, overcome the exit lens barrier potential, reach q_{eject} and be ejected (March, 1997). Ejection of ions can be performed to transmit ions to detectors or to remove unwanted ions. Selected methods of ion trapping, isolation and ejection for Thermo's LTQ and AB SCIEX's LITs are illustrated in Fig 1.11D.

A major limitation of the LIT is the so-called low-mass cut-off (LMCO). To fragment a peptide, the LIT first isolates it then, with resonant excitation, accelerates the ion through inert gas to cause fragmentation. However, due to this additional RF voltage, low-mass fragment ions (<28% of the precursor *m/z*) may have an initial q_z greater than q_{eject} and won't be trapped (Douglas et al., 2005; March, 1997). Overall, LITs are low-resolution instruments, operate at a constant bandwidth, fast scan speeds, and have a mass accuracy around 100 ppm (Watson and Sparkman, 2007).

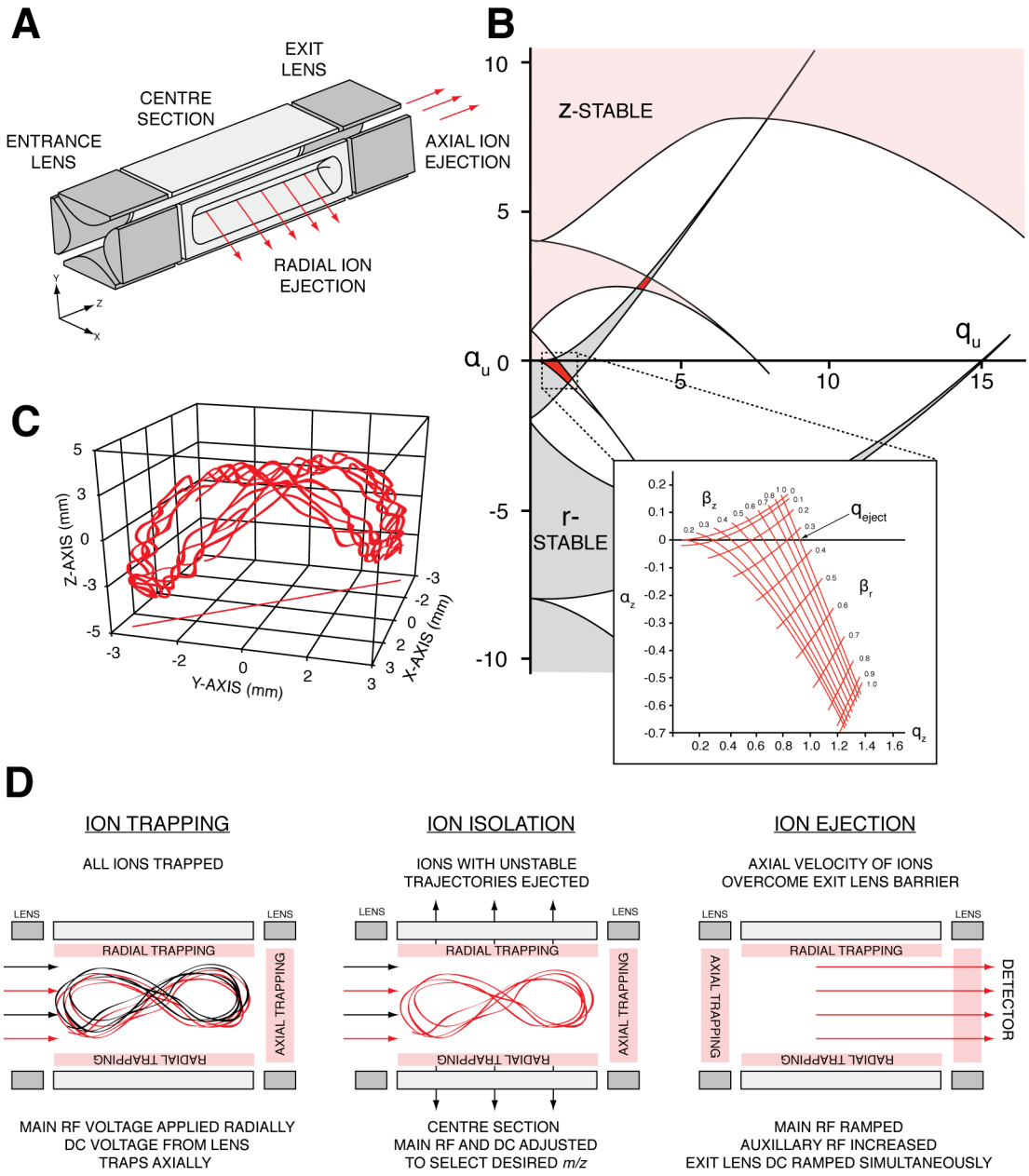


Fig. 1.11. Ion trap mass spectrometry. (A) Schematic of a linear ion trap (LIT) and different modes of ion ejection. Based on Schwartz et al., 2002. (B) Mathieu stability diagram for a quadrupole ion trap. Regions of simultaneous overlap of r- and z-directions are highlighted in red. Zoom of stability region with q_{eject} or low-mass cut-off indicated (q_z = 0.908) indicated. Based on March, 1997 (C) Shape of a stable ion's trajectory stored in a 3D ion trap. Based on Nappi et al., 1997. (D) Mechanisms of ion trapping, ion isolation and axial ion ejection.

1.3.3D ION GUIDES AND COLLISION CELLS

Quadrupoles can act as ion guides or collision cells when operated in RF-only mode. Without DC voltage, the quadrupole will transmit a broad range of *m/z* values, as ions will be confined radially within the centre of the quadrupole. Ion guides are often employed immediately after the ion source, in the relatively high-pressure of the first-pumped region, to collisional cool and radially confine incoming ions that reduces ion spread and enhances ion transmission (Douglas, 2009; Watson and Sparkman, 2007). When operated as a collision cell, the quadrupole accelerates ions through inert gas (usually argon or nitrogen) using a voltage between 20 and 200 eV to fragment ions, in a process termed collision-induced dissociation (CID); discussed further in section 3.5. Following fragmentation, the collision gas collisionally cools the fragment ions enabling refocusing of the ion beam and effective transmission out of the cell.

Other multipole configurations, e.g. hexapoles and octapoles, are also employed as ion guides and collision cells and transmit wider ion bandwidths than quadrupoles due to their steeper potential wells that lie closer to the rods. However, they cannot be employed as mass filters because their electric fields in x and y are coupled, meaning they cannot exploit stability regions (Douglas, 2009).

1.3.3E ORBITRAP MASS ANALYSER

The Orbitrap is the latest development in ion trapping and mass analyser technology, transforming the field since its commercialisation in 2005. This device, not too much wider than a two-pence coin, is a derivation of the Kingdon trap (Kingdon, 1923), the first device capable of storing ions in an electrostatic field. The Kingdon trap comprised of a thin wire, acting as a cathode, surrounded by a cylindrical anode with two end-cap electrodes that have DC current applied. Knight further developed the Kingdon trap by changing the shape of the outer electrode and separated it into two sections (Knight, 1981). This added an axial electric field enabling ions to harmonically oscillate in the z-direction. Neither the Kingdon nor Knight trap was capable of generating a mass spectrum.

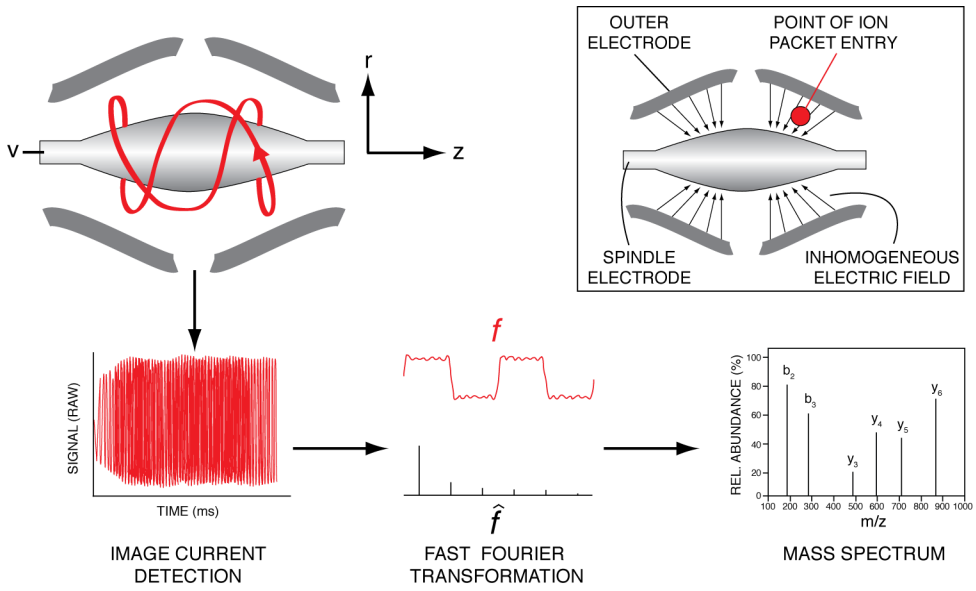


Fig. 1.12. Orbitrap mass spectrometry. The Orbitrap mass analyser consists of an outer electrode and spindle electrode that together generate an inhomogeneous, symmetrical electrostatic field (inset). When trapped, ions demonstrate harmonic oscillations along the z-axis at frequencies dependent on their mass. These oscillations are detected in the barrel electrode as a time-domain signal using image current detection, which is converted into mass spectra using fast Fourier transformation. Red arrow indicates the trajectory of a stable ion in the Orbitrap. Based on Makarov, 2000 and Watson and Sparkman, 2007.

The Orbitrap, developed by Makarov (Makarov, 2000), is a Knight-style Kingdon trap that orbitally confines ions in an electrostatic field and comprises of two coaxial axisymmetric electrodes: an outer barrel-like electrode and a central spindle-like electrode (Fig. 1.12). The outer barrel electrode is held at a near-ground potential whereas the spindle electrode has a constant DC potential of approximately -3400 V. Unlike the quadrupole ion trap, no oscillating electric fields are typically used.

When ions are injected directly into an electrostatic field they are not efficiently trapped, in fact they simply pass through. As a result, the ion injection process into the Orbitrap is complex: ions are first tightly focused in space and time in a curved linear trap (the C-trap), and injected into the Orbitrap, at a point offset from the centre, just as the Orbitrap's electric field is briefly increased (> 300 nsec). Without this careful synchronisation, efficient trapping (~30%) would not be achieved (Watson and Sparkman, 2007). The Orbitrap can trap all ions, irrespective of *m/z*, as all ion orbits have the same amplitude in the r-direction (Makarov, 2000). Once injected into the Orbitrap, ions experience a constant electrostatic field that accelerates them towards the middle of the trap, and then decelerates them as they reach the opposite end, where they stop and are accelerated back towards the

middle of the trap (de Hoffmann and Stroobant, 2013; Makarov, 2000; Watson and Sparkman, 2007). Ions therefore display harmonic oscillation around the spindle electrode along the z-axis (Fig. 1.12), analogous to a swinging pendulum bob. This occurs because the electric field is inhomogeneous but symmetrical due to the opposing surfaces of the axisymmetric coaxial electrodes being non-parallel. The field is inhomogeneous in two ways; first, the electric field is a gradient, which is at its weakest in the middle of the trap and increases uniformly in opposing directions along the z-axis; second, the vectors of the electric field are non-parallel. It is this second aspect that causes ions to oscillate along the z-axis at a frequency dependent on their mass and enables the Orbitrap to function as a mass spectrometer. These oscillations are detected in the barrel-like electrode, by two electrically isolated components, as a time-domain signal using image current detection. The image currents are then transformed into mass spectra using fast Fourier transform, similar to the method used in FTICR MS (de Hoffmann and Stroobant, 2013; Makarov, 2000).

Overall, the Orbitrap is a high-resolution device (100,000+), capable of sub ppm mass accuracy, a dynamic range of > 4,000, a mass range of 200-4000 *m/z* and simultaneous measurement of thousands of *m/z* values (Makarov et al., 2006a; Thermo, 2008; Watson and Sparkman, 2007).

1.3.4 HYBRID INSTRUMENTS

Most modern mass spectrometers rarely comprise of a single mass analyser but are constructed of multiple analysers of different types. These hybrid instruments can take advantage of each analyser's characteristics and therefore provide unique and/ or improved capabilities compared to the individual mass analysers alone. Examples of hybrid instruments are the AB SCIEX 5800 (TOF-TOF), AB SCIEX 5600+ (Q-TOF), Thermo Q Exactive (Q-collision cell-Orbitrap) and Waters Xevo TQ MS (QMF-collision cell-QMF). These are just a small selection of the types of mass spectrometer available today; however, each employs different analysers in different configurations meaning they have specific capabilities that will be suited for certain tasks and possibly not others. Below, the hybrid mass spectrometers employed in this thesis are discussed: the 4000 QTRAP and LTQ Orbitrap XL.

1.3.4A 4000 QTRAP

The 4000 QTRAP® is a triple quadrupole (QqQ) mass spectrometer, introduced by Applied Biosystems/ MDS Sciex (now AB SCIEX) in 2003 and is very similar in design to the original QqQ developed in 1979 (Yost and Enke, 1979). The 4000 QTRAP comprises of three quadrupoles: a QMF (Q1), a collision cell (q2) and an LIT (Q3) that has dual function as either a LIT or a QMF (Fig. 1.13A) (ABSCIEX, 2010; Hager, 2002). In this thesis the 4000 QTRAP was used predominately in two modes: selective-reaction monitoring (SRM) (Anderson and Hunter, 2006) and multiple reaction monitoring-initiated detection and sequencing (MIDAS) (Unwin et al., 2005).

In SRM, also referred to as multiple-reaction monitoring (MRM), a precursor ion is isolated, fragmented and the intensity of several selected fragment ions measured consecutively (Fig. 1.13B). The QTRAP system performs SRM scans by collimating and collisionally cooling an unresolved peptide ion beam in the Q0 ion guide, before it is transmitted into Q1 QMF that allows ions of a single *m/z* value to experience a stable trajectory and be transmitted into q2, the collision cell. Here selected precursor ions are rapidly accelerated through inert nitrogen gas, causing them to fragment, before continuing into Q3. Operating in QMF mode, Q3 transmits a specific fragment ion to the EM detector where signal intensity is recorded. The selected precursor and fragment (product) ion masses can be altered every few msec, allowing multiple precursor/ fragment ion pairs, known as transitions, to be monitored in rapid succession. This results in a highly selective and sensitive approach for peptide quantification; further discussed in section 3.6.2. In MIDAS, SRM scans are performed for user-defined transitions similarly to that described above. However, when the signal intensity for a transition reaches a user-defined threshold, Q3 switches from QMF to LIT mode. Since fragmentation occurs in the collision cell, all fragment ions are trapped without LMCO and accumulated in the LIT, before axial ejection to the detector to acquire an MS/MS spectrum. MIDAS is particularly useful for identification of PTM sites and for validation of SRM transitions amongst other applications (Lange et al., 2008; Unwin et al., 2009). Other scan modes available on the 4000 QTRAP include neutral loss scan, precursor ion scan and information dependent acquisition (IDA).

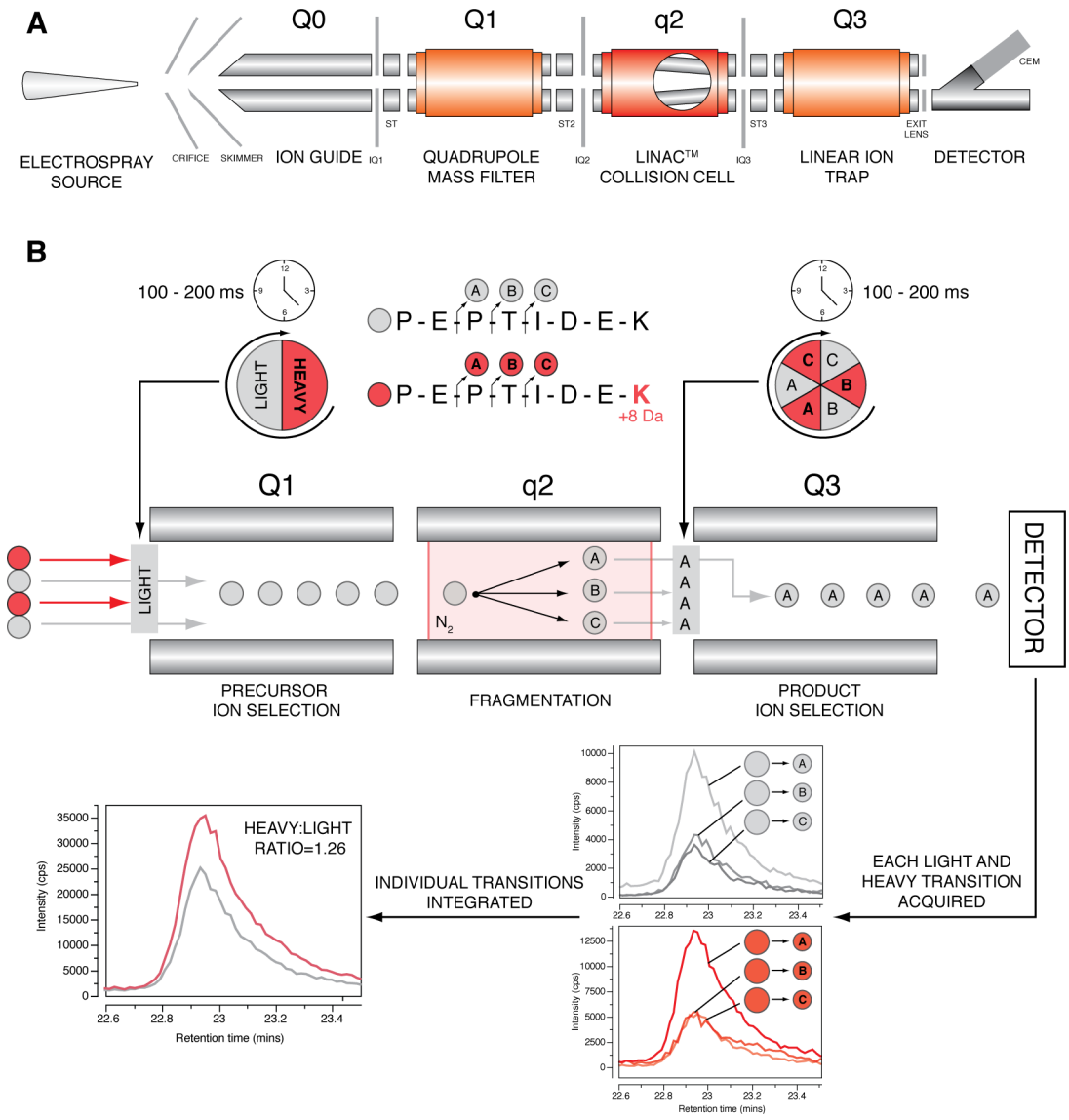


Fig. 1.13. The 4000 QTRAP® triple quadrupole mass spectrometer (AB SCIEX) and operating principles of Selected Reaction Monitoring (SRM). (A) Schematic of the 4000 QTRAP. It comprises of an ion guide, for collimating the ion beam and enhancing ion transmission; a quadrupole mass filter (QMF), which filters sample ions according to *m/z* value; a linear accelerator (LINAC™) collision cell, for fragmentation of selected sample ions (cut out shows how its rods are angled); and a linear ion trap that can also function as a QMF. Based on ABSCIEX, 2010 and Hager, 2002. (B) Overview of SRM for peptide quantification. Q1 QMF alternates transmission of light (endogenous) and heavy (standard) versions of the same peptide into the q2 collision cell for fragmentation, before being sent into Q3. Operating as a QMF, Q3 alternates transmission of predetermined fragment ions to the detector where signal intensity is measured. Multiple precursor and fragment ion pairs (transitions) are monitored for both heavy and light forms of the peptide, using identical dwell times, to measure a proportion of the peptide's ion current. When the transition signals are integrated, absolute quantification of the light (endogenous) peptide can be performed, as the abundance of heavy (standard) peptide is known.

1.3.4B LTQ ORBITRAP XL

Introduced in 2007, the LTQ Orbitrap XL is the second major release of Thermo's Orbitrap family. This instrument comprises of: quadrupole and octapole ion guides, a linear ion trap, C-trap, HCD collision cell and an Orbitrap mass analyser (Fig. 1.14A) (Makarov et al., 2006a; Makarov et al., 2006b; Thermo, 2008).

In the present thesis, the LTQ Orbitrap XL was used exclusively in data-dependent acquisition (DDA) mode, which enables the instrument to combine the Orbitrap's high-resolution scanning capabilities with the LTQ's rapid fragmentation and scan speeds. In DDA, three multipole ion guides transmit peptide ions through stages of differential pumping towards the LTQ. In the LTQ, a pre-determined number of ions are trapped before they are axially ejected into a region of higher vacuum through an octapole and into the C-trap. Here the ion packet is focused and injected into the Orbitrap that performs a high-resolution MS1 scan, providing precursor ion *m/z* values with high mass accuracy. A user-defined number of the most abundant ions (six was employed in this thesis) are then reported to the LTQ, which sequentially isolates, fragments and radially ejects the subsequent fragment ions to EM detectors (Fig. 1.14B and C). To fragment ions, the LTQ uses resonance excitation to increase the precursor ion velocity and therefore the energy of collisions between precursor ions and inert helium buffer gas (discussed further in section 3.5.1); as a result the LTQ exhibits LMCO. The LTQ and Orbitrap function in parallel, so that as the Orbitrap is acquiring an MS1 spectrum, the LTQ is isolating and fragmenting ions reported from the previous MS1 scan. Overall, this process of measuring the *m/z* of precursor then subsequent fragment ions is termed tandem MS or MS/MS and summarised in Fig 1.14B.

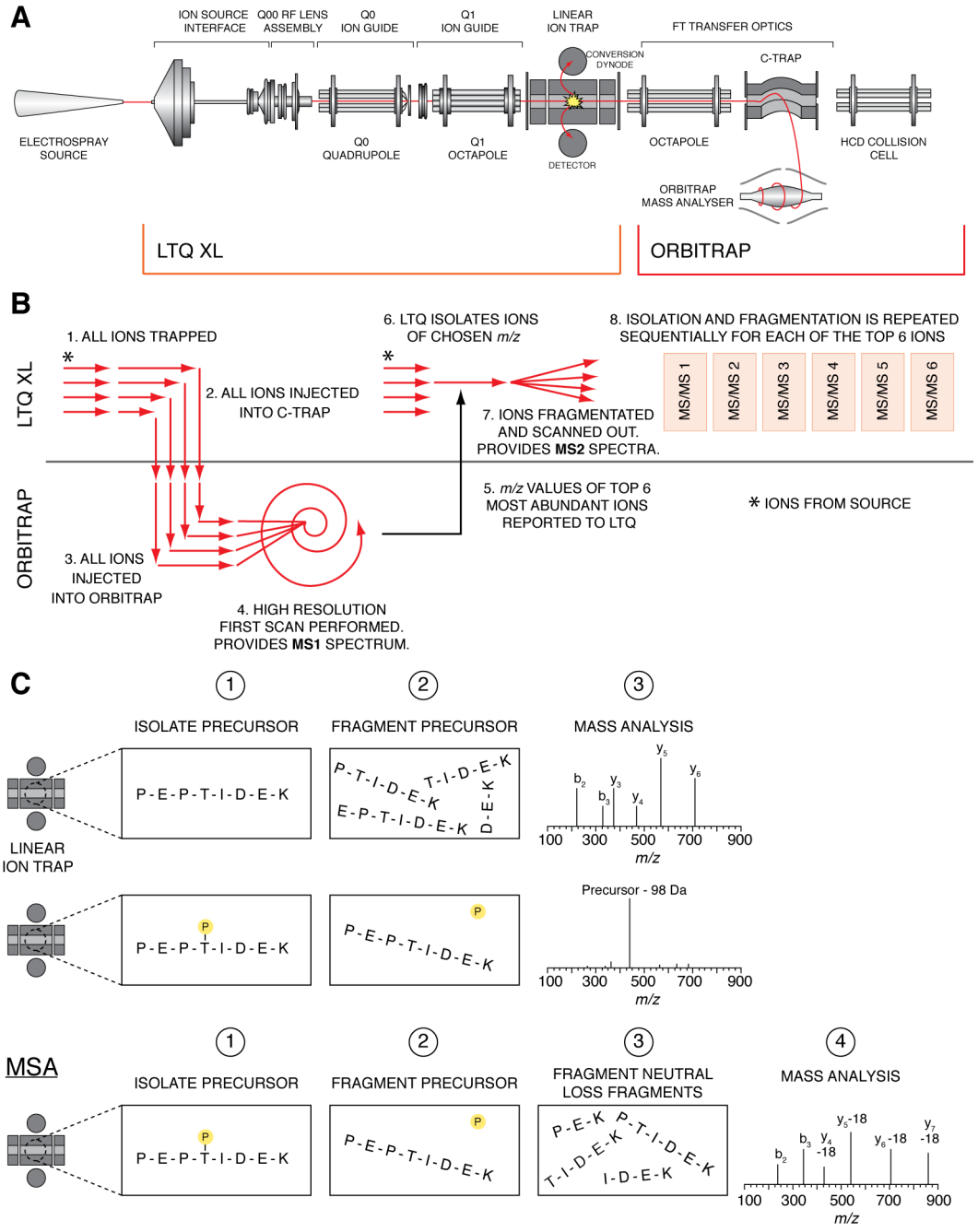


Fig. 1.14. The LTQ Orbitrap XL™ (Thermo Scientific). (A) Schematic of the mass spectrometer. Red line indicates the ion beam. Based on Thermo, 2008. (B) Data-dependent, parallel acquisition of the Orbitrap and linear trap quadrupole (LTQ). Following ionisation, peptide ions are trapped by the LTQ before being guided into the C-trap, which tightly focuses and efficiently transfers them into the Orbitrap where a high-resolution (≥ 30,000) scan is performed, generating an MS1 spectrum. *m/z* values of the top 6 precursor ions from the MS1 are reported to the LTQ where they are sequentially isolated, fragmented and product ions scanned out to generate MS/MS spectra. (C) Peptide fragmentation techniques in the LTQ. Collision-induced dissociation (CID) commonly provides structurally informative product ions (top). However, when a post-translational modification is present (e.g. phosphorylation), fragmentation can be poor making peptide identification difficult (middle). Multistage activation (MSA) employs additional activation at the *m/z(s)* where, in theory, the neutral loss ion(s) should reside (bottom) (Schroeder et al., 2004).

1.3.5 PEPTIDE FRAGMENTATION

MS-based proteomics relies on peptide fragmentation to generate information on peptide sequences and sites of post-translational modifications. There are different fragmentation techniques available in MS; those most commonly used are collision-induced dissociation (CID) (Haddon and McLafferty, 1968; Jennings, 1968; McLafferty and Bryce, 1967), electron capture dissociation (ECD) (Zubarev et al., 1998) and electron transfer dissociation (ETD) (Syka et al., 2004). In CID, precursor ions are accelerated to higher kinetic energy and allowed to collide with inert gas molecules to cause fragmentation. In ECD, peptide ions are irradiated with low-energy electrons to cause random backbone fragmentation through radical-driven pathways. Because ECD has essentially been limited to FTICR instruments, ETD, an ion-ion analogue of ECD compatible with ion traps and multipoles, was developed. In ETD, electrons are transferred to peptide ions from a radical anion, e.g. anthracene and fluoranthene, to fragment the peptide backbone. CID, the sole dissociation technique employed in this thesis, is currently the most popular and established means of gas phase fragmentation of peptides in proteomics and is discussed further below.

1.3.5A COLLISION-INDUCED DISSOCIATION OF PROTONATED PEPTIDES

In CID, peptide ions are accelerated by electric potentials and allowed to collide with inert neutral gas, e.g. nitrogen or helium. These collisions partially convert an ion's kinetic energy into internal vibrational energy that is distributed throughout the molecule. According to the 'mobile proton' model, first introduced by the groups of Wysocki (Dongre et al., 1996) and Gaskell (Burlet et al., 1992), this causes additional proton(s) to mobilise and populate energetically less favourable protonation sites. Being multifunctional compounds, peptides have multiple protonation sites, e.g. terminal amino groups, amide oxygens and nitrogens, and side chain groups. Protonation of amide nitrogens is favourable for peptide decomposition, as this weakens the amide bond and renders the neighbouring carbonyl group susceptible to nucleophilic attack by next neighbouring carbonyl groups or nearby electron-rich groups (Paizs and Suhai, 2004). Because different protonated species of the same peptide can coexist, fragmentation occurs at various sites along the peptide backbone, usually at amide bonds, thus generating

sequence-informative b- and y-ions, as well as other ions, e.g. a-ions, internal ions, immonium ions and non-sequence ions related to water/ ammonia loss (Fig. 1.15A). As the mobile proton model only provides a qualitative understanding of peptide fragmentation, efforts have been made to explain fragmentation in a semi-quantitative manner by predicting ion intensity relationships in MS/MS spectra. The emerging 'pathways in competition' (PIC) model considers all known fragmentation pathways, including pre-dissociation, dissociation and post-dissociation events (Paizs and Suhai, 2005).

Low-energy CID, the technique used in this thesis, employs a collision energy range of 1-100 eV, pressures of ~10 mTorr and activation times 10-30 msec (Wells and McLuckey, 2005). It fragments peptides by breaking their weakest bonds, predominantly through charge-directed rearrangement reactions, e.g. the b_x-y_z and 'diketopiperazine' fragmentation pathways. In the b_x-y_z pathway (Paizs and Suhai, 2001) (Fig. 1.15B), an additional proton mobilises to the nitrogen of the amide bond to be cleaved, before nucleophilic attack on the neighbouring carbonyl group from oxygen of the next neighbouring carbonyl group, thus forming b-ions with cyclic oxazolone structure and linear y-ions. A proton dimer (a short-lived, loose complex) may form between the fragment ions, which will dissociate following transfer of the additional proton to the ion with greatest proton affinity (Harrison, 1998; Paizs and Suhai, 2005). Diketopiperazine fragmentation pathways are also initiated by mobilisation of additional proton(s) to amide nitrogens (Cordero et al., 1993) (Fig. 1.15B), before nucleophilic attack by the terminal amino group on the neighbouring carbonyl group. Following amide bond cleavage, b-ions often dissociate as a neutral species, because they are unable to compete for the additional proton as large cyclic b-ions have much lower proton affinities than linear y-ions (Nold et al., 1999). As the mechanism of bond cleavage in the 'diketopiperazine' pathways differs depending upon which amide bond is targeted, it is referred to as diketopiperazine-y_{N-n}, where N is the total number of residues in the peptide and *n* the amide bond (aa(n)-aa(n+1)) to be cleaved. For example, the diketopiperazine-y_{N-2} pathway requires trans-cis isomerisation of the N-terminal amide bond to enable fragmentation, whereas this isn't required in the diketopiperazine-y_{N-n} (n>2) pathway (Paizs and Suhai, 2004). The above-mentioned fragmentation pathways are by no means an exhaustive list, as peptides may fragment through various other

mechanisms, e.g. the a_1-y_x pathway, the aziridinone pathway and nucleophilic attack by amino acid side chains (Paizs and Suhai, 2005).

Peptide fragmentation under low-energy CID conditions is not random and the eventual fragmentation pattern observed in MS/MS spectra depends on several factors, e.g. amino acid composition, instrument type and fragmentation technique. Peptide ions can be placed into two broad categories, the first category contains ions that have one or more highly favoured protonation sites, e.g. a singularly charged peptide with an arginine/lysine residue or a multiple charged peptide with many basic residues. This category is characterised by the large amount of energy required to mobilise the additional proton(s) to other protonation sites and will demonstrate limited fragmentation or resist fragmentation altogether. The second category contains peptides that have many protonation sites accessible within a narrow energy range and therefore suited for MS/MS analysis with CID. This category is best represented by doubly charged tryptic peptide (Paizs and Suhai, 2005). Numerous studies have shown residue-specific cleavage, especially at proline, aspartic acid and glutamic acid residues. Proline-containing peptides exhibit a distinct, highly abundant y-ion N-terminal of proline (Schwartz and Bursey, 1992). This has been attributed to proline's high proton affinity and also that b-ion formation is not favoured C-terminal to this residue (Vaisar and Urban, 1996). Glutamic acid- and aspartic acid-containing peptides generate abundant b-ions C-terminal to these residues. Peptides preferentially fragment at these points possibly because the bond is more labile than others (Qin and Chait, 1995; Yu et al., 1993).

The instrument type used for fragmentation can also affect which product ions are detected. LITs and collision cells are unlikely to generate identical MS/MS spectra, despite both utilising low-energy CID. Because LITs use resonant excitation CID, fragment ions aren't exposed to collision energy once produced, allowing labile fragment ions (e.g. large b ions) to remain stable and be detected when scanned out. Whereas collision cells expose all ions to collision energy, causing labile product ions to further fragment and not be detected. Additionally, collision cells have no LMCO enabling detection of low mass fragment ions, while LITs suffer from LMCO and are unable to trap such ions during fragmentation.

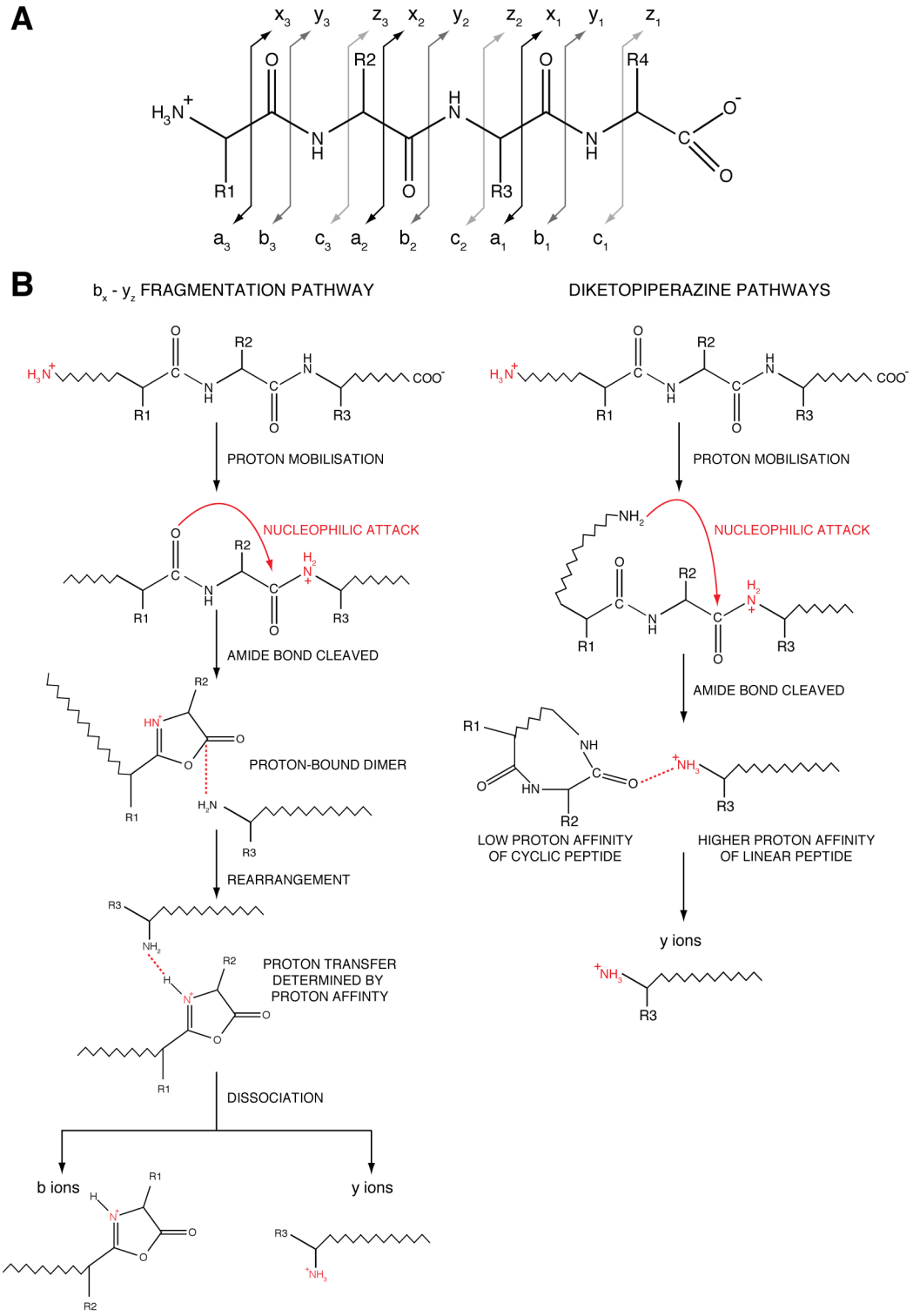


Fig. 1.15. Fragmentation of protonated peptides. (A) Nomenclature developed by (Roepstorff and Fohlman, 1984) and adapted by (Biemann, 1988). (B) Charge-directed fragmentation pathways common in low-energy collision-induced dissociation. Following collisions between peptide ions and inert molecules in a collision cell/ ion trap, the internal energy of the ions increases causing their extra proton(s) to mobilise and populate energetically less favourable sites. Protonation of amide nitrogens considerably weakens amide bonds and permits nucleophilic attack by functional groups: the N-terminal adjacent amide oxygen (b_x – y_z pathway, left) or the nitrogen from the N-terminal amino group (diketopiperazine pathways, right). Based on (Paizs and Suhai, 2005).

1.3.6 QUANTIFICATION STRATEGIES IN MASS SPECTROMETRY-BASED PROTEOMICS

Quantitative proteomics' ability to measure the abundance of thousands of proteins in many biological systems has seen it become a widely applied approach. Especially since mRNA levels cannot accurately predict protein expression levels (Gygi et al., 1999b) and that these protein levels reflect the cell's ultimate decision on its physiological needs. Quantifying the differences in a biological system between two or more physiological states is not only one the most important tasks in proteomics, but it is also one of the most challenging. Many MS-based quantification methods utilise differential stable isotope labelling to create specific mass tags, which are recognised by the mass spectrometer and enable quantification by comparing their respective signal intensities. These mass tags can be introduced metabolically, chemically, enzymatically or through spiked synthetic peptides/ proteins (Bantscheff et al., 2007). As a result of constantly improving LC-MS equipment, label-free quantification strategies, which forgo an isotope label, have grown in popularity and perform quantification by comparing samples after independent analyses. The labelling techniques employed in this thesis, stable isotope labelling by amino acids in cell culture (SILAC) (Ong et al., 2002) and protein standard absolute quantification (PSAQ) (Dupuis et al., 2008), are discussed below, along with alternatives. A summary of various MS-based quantification techniques is provided in Fig 1.9.

1.3.6A SHOTGUN PROTEOMICS AND RELATIVE QUANTIFICATION

Shotgun proteomics refers to the use of bottom-up proteomic techniques, described in section 3.2, in identifying peptides in complex mixtures. Named after shotgun DNA sequencing, where long DNA sequences are reconstructed computationally from many short reads, shotgun proteomics identifies peptides by matching MS/MS spectra of proteotypic peptides with those in a database (discussed in section 3.7) enabling identification of the protein (Washburn et al., 2001). Heavily used in discovery proteomics today, this approach has become a powerful tool in the unbiased analysis of protein abundance, modifications and interactions. However, coupling accurate peptide quantification with shotgun proteomics has been challenging, because peptides have unique physicochemical properties, e.g.

hydrophobicity, charge, size, amino acid composition, they will be unequal in their mass spectrometric response, even if they are from the same protein (Bantscheff et al., 2007). Nevertheless, by comparing each individual peptide between experiments, accurate quantification can be achieved. This was realised following the introduction of stable-isotope labelling of peptides in 1999 (Gygi et al., 1999a; Oda et al., 1999; Paša-Tolić et al., 1999). Since an isotope-labelled peptide is chemically identical to its non-labelled counterpart, they are expected to behave identically during MS analysis, yet the mass spectrometer can differentiate between the two peptide forms and quantification is achieved through comparison of their signal intensities.

Metabolic labelling provides the earliest possible point to introduce a stable isotope signature into proteins, thus minimising systematic errors, and was pioneered using ^{15}N-labelling of bacteria (Oda et al., 1999). Presently, SILAC is one of most popular forms of metabolic labelling (Ong et al., 2002), which in its original configuration utilised leucine isotopes but now more commonly uses isotopes of arginine and lysine, such that tryptic peptides will carry at least one label (Fig. 1.16A). In a triplex SILAC experiment, cells representing different conditions are grown in 'light', 'medium' or 'heavy' media. All amino acids in 'light' media contain ^{12}C and ^{14}N (i.e. no labelling), whereas 'medium' and 'heavy' are depleted of unlabelled arginine and lysine. Instead, 'medium' contains 4,4,5,5-D_4 L-lysine (+4 Da) and $^{13}C_6$ L-arginine (+6 Da), and 'heavy' $^{13}C_6$ $^{15}N_2$ L-lysine (+8 Da) and $^{13}C_6$ $^{15}N_4$ L-arginine (+10 Da) isotopes. Before MS analysis, the cells are lysed, lysates mixed at a 1:1:1 protein ratio, then, depending on the aim of the experiment, subject to selected sample preparation techniques. Due to the differential labelling, lysine-containing peptides from the 'medium' and 'heavy' conditions can be recognised by their +4 and +8 Da shift in the MS1 spectrum, respectively. Similarly, arginine-containing peptides will exhibit a +6 or +10 Da shift. Comparing the extracted ion chromatograms (XICs) integrated from the MS1 signal intensities of the SILAC triplet enables relative quantification, whilst MS2 spectra provide information for peptide identification. This approach enables the simultaneous identification and quantification of thousands of peptides in various biological systems, provided that the system can be appropriately labelled with stable isotopes. It is reported to have a linear dynamic range of between 1 and 2 orders of magnitude (Bantscheff et al., 2007).

Isotope signatures can also be introduced following protein isolation, as is the case for the isotope-coded affinity tags (ICAT) technique. Here proteins are chemically labelled by targeting cysteine thiol groups through covalent attachment of 'light' (containing no deuterium) or 'heavy' (containing 8 deuteriums) reagents prior to digestion (Gygi et al., 1999a). 'Light' and 'heavy' peptides are then distinguished by a +8 Da shift in MS1 spectra. However, deuterium causes a chromatographic shift that produces 'doublet' peaks, which correspond to the deuterium-labelled and unlabelled peptide. Other quantification techniques target digested peptides for labelling. In dimethyl labelling, formaldehyde with cyanoborohydride is covalently bound to digested peptides (Hsu et al., 2003), which contain either no label ('light', +28 Da) deuterium and/or ^{13}C ('medium', +32 Da; 'heavy', +36 Da). In isobaric tags for relative and absolute quantitation (iTRAQ, Fig. 1.16B), isobaric mass labels are covalently attached to peptide N-termini or lysine residues (Ross et al., 2004). The chromatographically indistinguishable isobaric tags generate low-mass reporter ions and therefore a MS/MS signature that supports quantification. Tandem mass tagging (Thompson et al., 2003), as well as the non-isobaric variant of iTRAQ, mTRAQ (DeSouza et al., 2008), adopt a somewhat similar approach. An additional approach to label peptides is to introduce a heavy oxygen isotope during digestion (Yao et al., 2001). Here two ^{18}O atoms are incorporated into the C-termini of peptides, resulting in a +4 Da shift in the MS1 spectra for labelled peptides. Label-free quantification, as the name implies, attempts to quantify proteins without the use of isotope labels, instead using spectrum counting- or MS1-intensity-based approaches. Reviewing the different label-free methodologies in-depth is beyond the scope of this thesis (for a recent excellent review refer to (Bantscheff et al., 2012)). To briefly mention a few examples, the number of peptide-to-spectrum matches or total ion currents (TICs) from MS/MS spectra have been used as a readout of protein abundance (Asara et al., 2008; Liu et al., 2004); plus integrating MS1 ion intensities of any given peptides over its chromatographic peak between experiments has shown to provide excellent linear responses (Bondarenko et al., 2002). One methodology of note is intensity-based absolute quantification (iBAQ) (Schwanhausser et al., 2011), where the intensities of observed peptides from any given protein are integrated and divided by the number of theoretically observable peptides to give an 'iBAQ intensity'. iBAQ intensities are then compared to those of a standard of 48 unlabelled proteins, spiked in at known concentrations, using linear regression, to provide an approximate abundance for thousands of proteins.

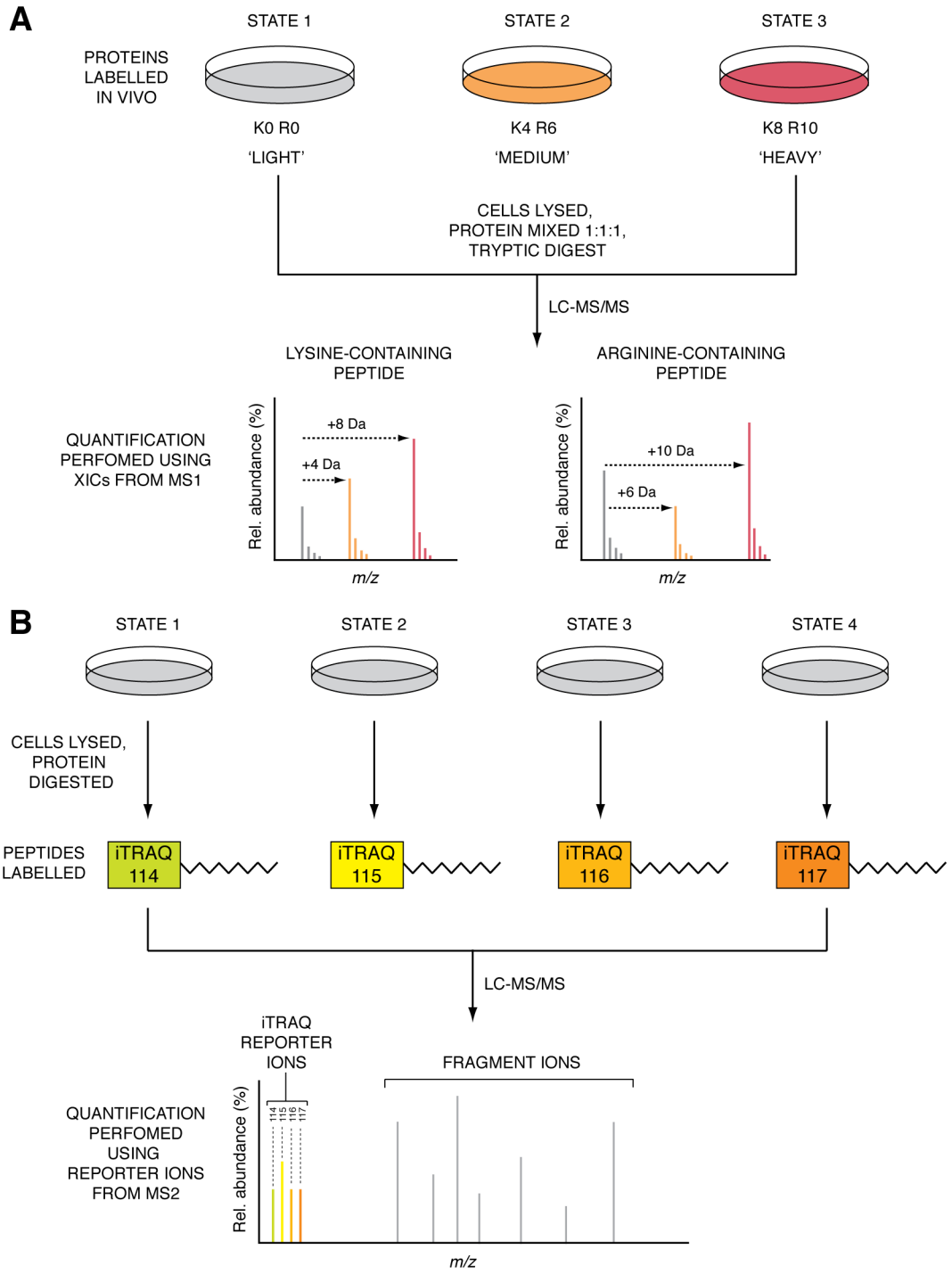


Fig. 1.16. Common strategies in quantitative proteomics using stable-isotope labelling. (A) Stable isotope labelling by amino acids in cell culture (SILAC) (Ong et al., 2002). Cells are metabolically labelled after being grown in media containing different isotopes of lysine and arginine. Cells are lysed, lysates mixed then subject to a chosen sample preparation technique, before LC-MS/MS. A characteristic shift in the MS1 spectra enables the peptides origin to be determined and relative quantification occurs through comparison of the integrated SILAC triplet extracted ion chromatograms (XICs) (B) Isobaric tags for relative and absolute quantitation (iTRAQ) (Ross et al., 2004). Isotope-coded tags are covalently bound to the digested peptides' N-termini or lysine residues. The tags are indistinguishable chromatographically but generate a reporter ion series in MS2 to enable quantification.

1.3.6B TARGETED PROTEOMICS AND ABSOLUTE QUANTIFICATION

Although shotgun proteomics is a powerful tool, it is not optimal for the systematic quantification of proteins in biological samples due to the usual high complexity of such samples, the limited sensitivity and stochastic nature of the approach and that, most often, relative quantification of the sample's most abundant proteins is performed. Consequently, to complement discovery-based strategies, techniques have been devised that target known peptides for accurate absolute quantification from complex matrices at low abundances.

SRM has been the reference quantitative technique for small molecules for several decades (Baty and Robinson, 1977; Zakett et al., 1978), but over the past few years it has emerged as an ideal tool for targeted proteomics because of its very high sensitivity (femtomole to attomole range), wide linear dynamic range (4 to 5 orders of magnitude) and high selectivity (Gallien et al., 2011; Stahl-Zeng et al., 2007). Targeted SRM approaches measure the signal intensity (peak area) of precursor/ fragment ion pairs (transitions) for a set of predefined peptides, over their chromatographic peaks, to achieve quantification (Anderson and Hunter, 2006). Such transitions are generally acquired using a QqQ, where precursor ions are selected in Q1, fragmented using CID in q2 and predefined fragment ions transmitted to detectors by Q3 (described in section 3.4.1, Fig. 1.13). Q1 is typically set to alternate transmission of precursor ions corresponding to unlabelled and stable-isotope-labelled versions of the same peptide for fragmentation in the collision cell. Coinciding with Q1, Q3 alternates transmission of (multiple) selected fragment ions to the detector, such that a large proportion of a peptide's total ion current (TIC) is measured (depending on amino acid composition, the top 5 transitions usually accounts for >70% of a peptide's TIC) (Lange et al., 2008). Signal-to-noise is low, as only the precursor of ion of interest is selected in Q1, and selectivity is high because of the requirement for both precursor and fragment ions being present. Consequently, when an unlabelled biological sample has an isotope-labelled standard spiked in (spike-in strategies discussed later), comparison between the relative intensities of the unlabelled and labelled transitions enables accurate and precise quantification. Furthermore, SRM assays can be segmented, such that peptides are monitored only when they are eluting from the LC column,

enabling a large number of peptides to be quantified in a single LC run (Stahl-Zeng et al., 2007).

Unlike unbiased shotgun proteomics, a minimum amount of prior information is required before an SRM assay can begin. This information is gathered during assay development, a slow and iterative process, that comprises of the following steps: (1) selection of target protein(s), (2) identification of proteotypic peptides, (3) selection of suitable transitions for each peptide, (4) optimisation of instrument parameters for each transition/peptide (e.g. collision energy, dwell times, retention times etc.), (5) determination of detection and quantification limit, (6) determination of peptide signal response across different concentrations, (7) validation of each transition (Gallien et al., 2011; Lange et al., 2008). Only once these steps have been performed, can an SRM assay begin.

Similar to relative-quantification strategies, there are various ways to introduce an isotope label into a targeted proteomics workflow. But a key difference here is that by knowing the amount of an isotope-labelled standard, it is possible to measure the absolute amount of a peptide and, by inference, the protein of interest in a sample. AQUA peptides (AQUA an acronym for absolute quantification) are synthetic isotope-labelled peptides that may be spiked into peptide digests to act as surrogate standards for endogenous peptides (Barr et al., 1996; Gerber et al., 2003). QconCAT involves the generation of a concatemer of isotope-labelled peptides that is spiked into a sample immediately before digestion (Beynon et al., 2005). The concatemer can contain peptides from multiple proteins and is synthesised such that proteolytic digestion releases the isotope-labelled peptide standards. Yet, these techniques do have some drawbacks, chiefly they assume either complete protein digestion (as is the case for AQUA) or equal digestion between the concatemer and endogenous protein (as for QconCAT), as well as equivalent losses and recoveries of endogenous and isotope-labelled peptides, assumptions that are very likely not true (Gallien et al., 2011; Lowenthal et al., 2014). Nevertheless, these methods will undoubtedly provide reasonable data, but for truly accurate measurements they are likely to fall short. PSAQ however is recognised as the most accurate method currently available to quantify a protein, as this involves spike in of full-length, isotope-labelled protein into an unlabelled sample (Dupuis et al., 2008). Consequently, digestion conditions for the endogenous and labelled proteins are

identical, the effects of protein structure and amino acid sequence on digestion are equalised and systematic errors are avoided as the protein standard is added at the beginning of the workflow. Other methodologies similar to PSAQ have been described, e.g. full-length expressed stable isotope-labeled proteins for quantification (FLEXIQuant) (Singh et al., 2009) and absolute SILAC (Hanke et al., 2008). Interestingly, Hanke et al. did not utilise SRM or a QqQ, but selected ion monitoring (SIM) on an LTQ-Orbitrap to perform peptide quantification.

Accurate and consistent quantification of proteins is an essential requirement for systems biology to precisely construct metabolic networks and cellular signalling pathways. The development of SRM-targeted quantification, coupled with the above-mentioned spike-in strategies, has great potential to provide this required information. Yet, the field of proteomics is still very much in development, emerging techniques, e.g. sequential window acquisition of all theoretical fragment-ion spectra (SWATH) (Gillet et al., 2012), potentially have the power to deliver SRM-like quality data, while reducing assay develop time and significantly increasing the number of quantified proteins in a single run. It is no longer in the realms of science fiction to suggest that quantitative proteomics may retire the Western blot for protein quantification.

1.3.7 SOFTWARE TOOLS

A major challenge in MS-based proteomics today is that of data handling and/or interpretation. Because a single shotgun proteomics experiment generates a quantity of data that makes manual interpretation unfeasible, software packages designed to computationally interpret tandem MS data have become invaluable. A variety of software suites are currently available, from both academic and commercial sources, which are used for quantification, peptide-to-spectrum mapping and the identification and localisation of post-translational modifications. This section briefly describes a number of such software suites that are commonly used by the proteomics community.

1.3.7A QUANTIFICATION TOOLS

Over the past few years the number of software packages available to analyse proteomics data has risen dramatically. Researchers now have a variety of software

tools to choose from to computationally process, interpret and quantify their discovery- and targeted-based proteomic data. In shotgun proteomics, a popular and freely available software tool is MaxQuant (Cox and Mann, 2008), which has been used in hundreds of studies since its initial release in 2008 (judging by citations). Following its integration with the Andromeda database search engine (discussed in next section), the MaxQuant suite of algorithms incorporates all steps needed in a computational proteomics platform. Once data is input into MaxQuant, it is processed with the 'quant' module, before MS/MS spectra are matched with database-derived peptide sequences with Andromeda, and the subsequent data from each is combined and subject to statistical analysis using the 'identify' module. The 'quant' module searches MS1 for isotope clusters and SILAC peptide pairs, and performs quantification by integrating XICs. The 'identify' module compiles the final output files and calculates various statistics, e.g. peptide posterior error probability (PEP) score. Recent releases of MaxQuant are compatible with a variety of stable-isotope relative quantification techniques, e.g. SILAC, TMT and iTRAQ, as well as label-free methods, e.g. iBAQ. Freely available software alternatives for quantification of proteomic data include: OpenMS (Kohlbacher et al., 2007), Viper (Monroe et al., 2007) and Census (Park et al., 2008). Commercial alternatives include: Proteome Discoverer (Thermo), Progenesis LC-MS (nonlinear dynamics, a Waters company) and Mascot distiller (Matrix Science Ltd).

In targeted proteomics, Skyline is a popular, open-source software package for SRM assay setup and quantitative analysis, designed by the MacCoss group (MacLean et al., 2010). It is compatible with all spike-in quantification techniques discussed earlier and quantifies peptides by integrating the XICs of multiple transitions. As SRM assay setup is a slow and iterative process, functionality to import public spectral libraries, e.g. PeptideAtlas (Desiere et al., 2006), predict peptide retention times (Escher et al., 2012) and streamline collision energy optimisation was added to Skyline to aid with this process.

1.3.7B SEARCH ENGINES

Database search algorithms (search engines), which match MS/MS spectra to peptide sequences from curated databases, are critical components of modern discovery-based proteomic workflows (Marcotte, 2007; Nesvizhskii et al., 2007). The

first major attempt to correlate fragment ion spectra with database-derived peptide sequences was performed by the group of John Yates III, who described the Sequest strategy (Eng et al., 1994). The algorithms of this strategy search a protein database for peptides that match the molecular weight of the experimentally observed peptide, before a cross-correlation function is used to match *m/z* values of predicted fragment ions to those observed experimentally. A score is then assigned to describe the nearness of the match. Currently, Mascot (Matrix Science Ltd) (Perkins et al., 1999) and Andromeda (Cox et al., 2011) are two commonly used search engines. Mascot compares experiment-derived MS/MS spectra to *in silico*-generated data in a similar way to Sequest, but incorporates a probabilistic-scoring model by calculating the probability that an observed match is a chance event. This provides a simple rule to judge whether a match is significant and guards against false positives. Similarly, Andromeda uses probability-based scoring. It has now replaced Mascot as the search engine integrated into the MaxQuant environment (Cox and Mann, 2008) and utilises a scoring function derived from the p-score, introduced for the interpretation of MS^3 spectra (Olsen and Mann, 2004). Andromeda separates the spectra into 100 Th bins, before dynamically match testing the peaks within each bin, in an intensity-prioritised manner. When compared, Mascot and Andromeda demonstrated a high level of correlation in terms of peptide identifications, with few peptides identified exclusively with either approach (Cox et al., 2011). The techniques employed by these search engines to identify and localise post-translational modifications are discussed in section 3.8.2.

1.3.7C DATABASES

A variety of publically available databases have been developed to provide the proteomics community with complete human, mouse, rat and E.coli protein sequences, amongst others. These databases play a crucial role in enabling the identification of peptides and proteins using MS/MS spectral searching. The International Protein Index (IPI) (Kersey et al., 2004), retired in 2011 due to advances in data organisation, offered non-redundant data sets on human, mouse, and rat proteomes by combining Swiss-Prot, TrEMBL, Ensembl and RefSeq databases. IPI users are now directed to use the Universal Protein Resource Knowledge Base (UniProtKB) (Apweiler et al., 2004; UniProt, 2011), which expertly curates, manually annotates and reviews the Swiss-Prot database. Other

databases, e.g. NCBI and EMBL EST divisions, are available for peptide-to-spectrum mapping, but discussing these is beyond the scope of this thesis.

1.3.8 PHOSPHOPROTEOMICS

Out of the > 200 different PTMs that can occur on proteins (Wold, 1981), none are as universally recognised as reversible phosphorylation of serine, threonine and tyrosine residues to control nearly all aspects of life (Cohen, 2002). Phosphorylation plays a critical role in essentially all cellular processes through regulation of protein activity, sub-cellular localisation, expression levels and degradation, as well as protein-protein interactions. The phosphorylation status within a cell is highly dynamic, controlled by the action of kinases and phosphatases, and reflects the cell's current physiological needs. It is unsurprising that aberrant phosphorylation has been associated with the onset and progression of many pathological conditions, e.g. various cancers (Hanahan and Weinberg, 2000), Parkinson's (Greggio, 2012) and Alzheimer's disease (Gong et al., 2006). Consequently, understanding phosphorylation at the molecular level of normal and diseased cellular states has become an area of intense research.

It is estimated that ~30% of all mammalian proteins are phosphorylated at any one time, > 500 serine, threonine and tyrosine kinases comprise the human kinome, and that ~700,000 potential phosphorylation sites are available to any given kinase (Hunter, 1998; Manning et al., 2002; Ubersax and Ferrell, 2007). Phosphorylation can occur on serine, threonine, tyrosine, histidine, arginine, lysine, cysteine, glutamic acid, and aspartic acid residues, but by far the most prevalent sites are the hydroxyl groups of serine, threonine and tyrosine (Hunter, 1991). Out of all phosphorylation events, pS accounts for ~90%, pT ~10% and pY ~0.05% (Hunter and Sefton, 1980). Despite pY being the least abundant phosphorylation type, possibly explaining why it was discovered (fortuitously) relatively late (Eckhart et al., 1979), it plays an essential role in receptor-mediated signal transduction, with its aromatic ring affording stronger binding with protein interaction partners compared to the aliphatic chains of pS and pT (Hunter, 1998).

MS has revolutionised the characterisation of protein phosphorylation, as the large-scale identification and quantitation of phosphopeptides is now feasible following recent advances in liquid chromatography, sample preparation techniques and MS

instrumentation. However, sample complexity and the low stoichiometry of phosphopeptides complicate the study of the phosphoproteome by MS (Rigbolt and Blagoev, 2012). Phosphopeptides have also been reported to suffer from lower ionisation efficiencies compared to unmodified peptides (Ishihama et al., 2007), but this is disputed (Steen et al., 2006). For effective MS analysis, these issues necessitate the use of enrichment strategies, e.g. strong cation exchange (SCX) (Beausoleil et al., 2004), TiO_2 chromatography (Larsen et al., 2005; Pinkse et al., 2004), immobilised metal affinity chromatography (IMAC) (Andersson and Porath, 1986; Ficarro et al., 2002) and SCX followed by IMAC or TiO_2 (Lemeer et al., 2008; Villen and Gygi, 2008) amongst others. Although pY-containing peptides may be enriched using these methods, its very low abundance necessitates the use of pull-down strategies with pY-specific antibodies to achieve a high level of enrichment (Rush et al., 2005).

The versatility of phosphoproteomic and quantitation strategies allows them to be applied in combination to a wide range of biological processes in various organisms. Quantitative phosphoproteomics, using SILAC or iTRAQ, has enabled the interrogation of downstream signalling from RTKs (Hammond et al., 2010; Morandell et al., 2008) and GPCRs (Christensen et al., 2010; Hoffert et al., 2012); as well as the characterisation of phosphoproteomic changes in EphB2- and ephrin-B1-controlled cell sorting (Jorgensen et al., 2009), in Salmonella-infected cells (Rogers et al., 2011), and following stem cell differentiation (Rigbolt et al., 2011a). The successful application of quantitative phosphoproteomics to probe signalling events in different organisms, e.g. *Saccharomyces cerevisiae* (Saleem et al., 2010; Schreiber et al., 2012), *Escherichia Coli* (Macek et al., 2008) and Drosophila melanogaster (Chang et al., 2008), further demonstrates the wide scope of its utility.

1.3.8A COLLISION-INDUCED DISSOCIATION OF PHOSPHORYLATED PEPTIDES

Post-translational modification of residues on a peptide can redirect the preferred site of cleavage during CID. Phosphorylation of serine and threonine residues is the best known ‘MS/MS labile’ chemical modification, but other PTMs, e.g. glycosylation, sulfonation and nitrosylation, also fall into this category (Chicooree et al., 2014; Mikesh et al., 2006). Understanding the fragmentation behaviour of

phosphopeptides has enabled MS to become a valuable tool for recognising protein phosphorylation and pinpointing phosphorylation sites.

Since the phosphate group of a phosphopeptide is relatively labile, it provides a low-energy pathway that competes with backbone fragmentation in CID. Consequently, a MS/MS spectrum of a phosphopeptide typically contains a dominant neutral loss peak that is 98 (or occasionally 80 Da) lower than the precursor mass, corresponding to the loss of H_3PO_4 (or HPO_3) (DeGnore and Qin, 1998; Liao et al., 1994). Sometimes the neutral loss peak dominates to such an extent that other, sequence-informative ions are not observed. Fragmentation patterns of phosphoserine (pS) and -threonine (pT)-containing peptides differ in CID (Fig. 1.17A). pS-containing peptides fragment in a simple, charge state-independent manner, by losing predominately H_3PO_4 to form dehydroalanine. pT-containing peptides display more complex, charge-dependent fragmentation behaviour, by predominantly losing H_3PO_4 to form dehydrobutyrine or losing HPO_3 through dephosphorylation (DeGnore and Qin, 1998). Consequently, observing neutral loss of H_3PO_4 is a common method to recognise protein phosphorylation at serine and threonine residues. Neutral loss of H_3PO_4 can occur via two mechanisms: β-elimination or charge-directed pathways (Fig. 1.17B) (Boersema et al., 2009). In β-elimination, the phosphate group abstracts hydrogen from the α-carbon of the phosphorylated residue resulting in loss of the phosphate group (DeGnore and Qin, 1998; Tholey et al., 1999). In charge-directed mechanisms, the phosphate group abstracts the mobile proton, leading to nucleophilic attack by neighbouring groups or E2-elimination reactions (Fig. 1.17B) (Boersema et al., 2009). Charge-remote and charge-directed mechanisms are likely to be in competition, with the charge-directed mechanism possibly the most abundant (Palumbo et al., 2008; Tholey et al., 1999).

Generally, LITs will exhibit a higher level of neutral loss compared to collision cells, because the latter expose all ions to collision energy meaning they will automatically activate neutral loss fragments, whereas LITs usually activate the precursor ion only, which allows neutral loss fragments to remain intact. As a result, methods have been developed for LITs to obtain more sequence information from phosphorylated peptides by employing a second stage of activation, e.g. data-dependent neutral loss-triggered MS^3 (Beausoleil et al., 2004) and multistage activation (MSA) (Schroeder et al., 2004). In MS^3, a survey scan is performed to determine precursor

m/z values, followed by an MS/MS scan. If a dominant neutral loss peak is detected, a third step is triggered where the newly released neutral-loss precursor ion is further isolated and activated. As the neutral loss fragment no longer has a lower-energy fragmentation pathway from the phosphate group available, more extensive backbone fragmentation can be achieved. However, any benefits provided by MS^3 come at the expense of speed. In MSA, also referred to as pseudo-MS^3, a survey scan is performed before a precursor ion is isolated and activated at its *m/z*. This is followed by activation at *m/z* values where, in theory, different neutral loss fragments reside, while all other fragments remain trapped (Fig. 1.14C). MSA is likely to provide richer sequence information than MS^3, because ions formed in the first activation event, which are discarded by MS^3 and possibly not formed in subsequent activations, are included in the MSA spectrum. MSA can be viewed as generating a composite spectrum of the MS^2 and MS^3. Additionally, MSA has a speed advantage over MS^3, as it doesn't require refilling of the ion trap. Only a few studies have directly compared the utility of MS^2, MS^3 and MSA in large-scale phosphorylation analyses (Ulintz et al., 2009; Villen et al., 2008). But they differ in their conclusions regarding the benefit provided by additional activation steps, possibly because the studies used different neutral loss settings and MS setup. Villén et al. concluded that the extra time spent on additional activation events is not worthwhile; unlike Ulintz et al. who found MSA to outperform both MS^2 and MS^3 by identifying more phosphosites and that MSA spectra had the highest number of identifiable peaks and the most information content.

Phosphotyrosine (pY)-containing peptides are generally stable under low-energy CID conditions, occasionally demonstrating neutral loss of HPO_3. H_3PO_4 cannot be lost by pY, because the aromatic ring prevents E2-elimination or S_N2-neighbouring group participation reactions and stabilises the C-O bond of pY's C-O-P structure, causing fragmentation at the weaker O-P bond (Boersema et al., 2009; Tholey et al., 1999). It is therefore challenging to detect pY in LITs. To address this issue, higher collision energies, e.g. one-tenth of precursor *m/z*, have been employed in collision cells and C-traps to improve the detection of pY by generating a diagnostic immonium ion at *m/z* 216.043 (Fig. 1.16A) (Olsen et al., 2007; Steen et al., 2001). Even if this immonium ion was generated in an LIT, it is likely it would be lost due to LMCO.

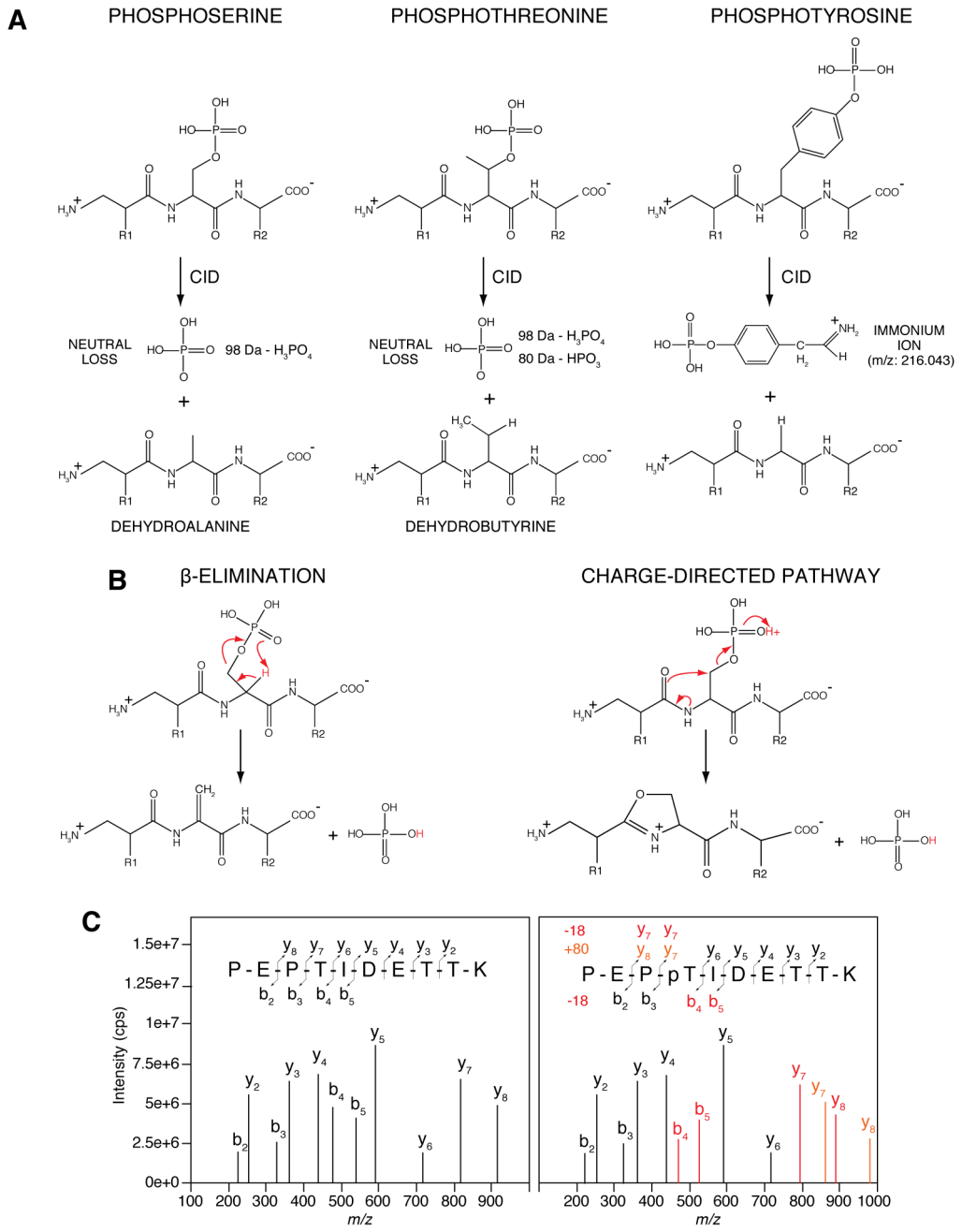


Fig. 1.17. Collision induced-dissociation of phosphopeptides. (A) Fragmentation of phosphoserine, -threonine and -tyrosine. During CID, the phosphomoiety is often lost in preference to peptide backbone fragmentation as a characteristic neutral loss. In contrast to phosphoserine, fragmentation of phosphothreonine and -tyrosine is complex. Phosphoserine predominately loses H_3PO_4; phosphothreonine typically loses H_3PO_4, but may also lose HPO_3; phosphotyrosine infrequently exhibits neutral loss but more often generates an immonium ion at higher collision energies. (B) Examples of neutral loss fragmentation pathways. β-elimination, the phosphate group abstracts hydrogen from the phosphorylated residue's α-carbon. Charge-directed pathway, protonation of the phosphate group leads to nucleophilic attack by the β-carbon's carbonyl group. Red arrows indicate movement of electron pairs; red hydrogen is that abstracted by the leaving phosphoric acid. Based on Boersema et al., 2009. (C) Determining phosphosite localisation. Left, an unmodified hypothetical peptide; right, same peptide but with T4 phosphorylated. Orange and red ions are 'site determining'. Orange, fragments with intact phosphate group attached (+80 Da from corresponding unmodified ion); red, fragments containing a dehydrobutyrine residue (-18 Da).

1.3.8B PHOSPHOSITE LOCALISATION

Pinpointing a phosphosite can be even more demanding than identifying the peptide itself, especially when considering ~16% of amino acids in an average human protein are serine (8.6%), threonine (5.4%) or tyrosine (2.3%) (Echols et al., 2002). To confidently localise a phosphosite, 'site determining' ions must be present in the MS/MS spectrum. To demonstrate how phosphosites are localised, Fig. 1.17C shows theoretical spectra of a peptide in phosphorylated and non-phosphorylated forms. The peptide has three potential phosphorylation sites: T4, T8 and T9. Compared to the non-phosphorylated peptide, fragment ions with an intact phosphate group will be +80 Da, while dehydroalanine/ dehydrobutyrine residues are -18 Da. T8 and T9 are not phosphorylated, as the y_6 ion and smaller y ions have identical *m/z* values as the unmodified peptide. T4 is likely to be phosphorylated as y_8 and y_7 ions are site determining, i.e. they are in +80 and -18 Da forms. Furthermore, b_5 and b_4 ions contain dehydrobutyrine residues, while b_3 and b_2 do not. For confident phosphosite localisation, fragment ions with an intact phosphate group are required, because loss of -18 Da can also correspond with the loss of water from the unmodified peptide (Boersema et al., 2009; Lehmann et al., 2007). When site-determining ions are lacking, phosphosite localisation could be prevented, especially if several possible phosphorylation sites exist on a peptide. Further complicating phosphosite localisation was the report of possible intramolecular transfer of phosphate between amino acids with a hydroxyl group during CID scrambling (Palumbo and Reid, 2008), but this has been disputed since (Aguiar et al., 2010).

Due to the sheer number of MS/MS spectra generated by a single phosphoproteomics experiment, manual inspection to identify phosphosite(s) has become extremely difficult, if not impossible. Consequently, in the interests of throughput and objectivity, data analysis software packages have incorporated various strategies to identify protein phosphorylation and possible phosphosites. Because reviewing these different strategies in detail is beyond the scope of this thesis, the methods employed by Mascot and MaxQuant/ Andromeda are taken as examples. Most strategies for scoring localisation reliability fall into two main categories; first, those that assess the probability of site determining ions to be matched at random; second, those that calculate a difference score between the

possible site localisations following a database search (Chalkley and Clauser, 2012). MaxQuant employs the former strategy with its PTM score (Olsen et al., 2006), while Mascot uses the latter with its delta score (Savitski et al., 2011). For peak picking, both MaxQuant and Mascot use a similar method by splitting MS/MS spectra into 100 or 110 Th bins, respectively, with the most intense peaks within each bin used for searching. For phosphosite localisation, the PTM score considers the 4 highest peaks per bin and tests all combinations of serine, threonine and tyrosine phosphorylation; those with the highest scores are then reported. In the Mascot delta score, following a database search, the difference between the best and second best Mascot ion score are calculated for alternative phosphosites on the same peptide. Both of these strategies are used heavily, as well as several others, e.g. A-score, SLoMo and the SLIP score, but a consensus is yet to be reached on a standard strategy for determining phosphosite localisation (Chalkley and Clauser, 2012).

1.4 OUTLINE AND OBJECTIVES OF THE THESIS

With modern biology adopting an ever-increasing holistic approach to study the behaviour of complex signalling pathways, through the use of mathematical and computational modelling, it is important that accurate measurements of pathway component concentrations are made. Although Ras proteins have been extensively studied for over 30 years, it is remains unclear how many Ras molecules are present in a cell, as well as which isoform is most abundant. While several attempts have been made to measure cellular Ras abundance in the past few years, the accuracy of the developed techniques is disputable and the results from these studies conflict. The present thesis aims to develop a selected reaction monitoring (SRM)-based assay, combined with the protein standard absolute quantification (PSAQ) technique, to enable the accurate, precise and reliable quantification of H-, K(A)-, K(B)- and N-Ras isoforms at endogenous levels. The Ras quantification technique should also be designed to be easily utilised by most biological laboratories, with access to a mass spectrometry facility, by using procedures that are recognisable to such laboratories and/ or do not require specialist knowledge. The target cell lines in this thesis are a panel of isogenic SW48 colorectal cell lines that harbour a range of heterozygous K-Ras mutations, as well as H- and N-Ras mutants, with each mutant sequence expressed under its respective endogenous promoter. These cell lines can reveal if Ras mutations can affect isoform protein expression levels.

Ras proto-oncogenes are commonly activated by mutations at codons 12, 13 and 61, which occur in ~16% of human cancers and each result in a constitutively active Ras protein. However, clinical data indicates that codon 12 and 13 mutant K-Ras proteins are unequal in their signalling output, since colorectal cancer patients harbouring these mutations display different overall survival rates and responsiveness to anti-EGFR antibody therapy. By subjecting isogenic SW48 cells harbouring G12D, G12V and G13D K-Ras mutations, which represent 75% of colorectal cancer cases that harbour a K-Ras mutation, to proteomic and phospho-proteomic analysis, using mass spectrometry-based techniques, this thesis aims to probe the differences in signalling output from the different K-Ras mutants.

2 – MATERIALS AND METHODS

2.1 GENERAL SOLUTIONS AND BUFFERS

All chemicals in this section and 2.2 were purchased from Sigma Aldrich (Poole, UK), unless stated otherwise.

RADIOIMMUNOPRECIPITATION ASSAY (RIPA) LYSIS BUFFER:
10 mM Tris HCl (pH 7.5), 150 mM NaCl, 0.1% [w/v] Triton X-100, 0.1% [w/v] SDS, 1% [w/v] sodium deoxycholate, supplemented with 50 mM NaF, mammalian protease inhibitors (diluted 1:250; #P8340, Sigma Aldrich), PhosSTOP tablet (1 per 10 mL; #04 906 845 001, Roche, West Sussex, UK), unless indicated otherwise.

NP40 LYSIS BUFFER:
25mM Tris HCl (pH 7.5), 0.5% [w/v] NP40, 100mM NaCl, supplemented with 50 mM NaF, mammalian protease inhibitors (diluted 1:250), PhosSTOP tablet (1 tablet per 10 mL).

TRANSFER BUFFER:
25 mM Tris-HCl (pH 8.5), 190 mM glycine, 10% [v/v] methanol.

10× TBS:
200 mM Tris HCl (pH 7.6), 1.5 M NaCl.

1× TBS-T:
0.1% [w/v] Tween 20 (#EC-407, National Diagnostics, Hull, UK) in 1× TBS.

BLOCKING SOLUTIONS:
1× TBS-T supplemented with 5% [w/v] non-fat dry milk powder (Marvel, Premier foods, Hertfordshire, UK), 3% [w/v] Gelatin from cold water fish skin (#G7765, Sigma Aldrich) or 5% [w/v] BSA.

DNA LOADING BUFFER:
5% [w/v] glycerol, 0.1 mM EDTA, ~0.005% [w/v] bromophenol blue.

5× SAMPLE BUFFER:

Tris HCl (pH 6.8), 15% [w/v] SDS, 50% [v/v] glycerol, 16% [v/v] beta-mercapto-ethanol, with bromophenol blue as a dye.

5× M9 MINIMAL MEDIA:

120 mM Na_2HPO_4, 55 mM KH_2PO_4, 21.5 mM NaCl, 1 mM $MgSO_4.7H_2O$, 0.5 mM $CaCl_2$, 48 mM NH_4Cl, 2% [w/v] glucose, 0.5 mg/ mL of each hydrophilic amino acid, 1 mg/ mL of each hydrophobic amino acid, not including arginine and lysine (pH 7.4).

2.2 SOLUTIONS AND BUFFERS FOR LC-MS/MS SAMPLE PREPARATION AND INSTRUMENTATION

AMBIC:

50 mM NH_4HCO_3 in HPLC-grade H_2O (#23595328, VWR, Leicestershire, UK).

REDUCTION BUFFER:

10 mM dithiothreitol (DTT) in 50 mM Ambic.

ALKYLATION BUFFER:

55 mM iodoacetamide (IAA) in 50 mM Ambic.

SDT BUFFER:

100 mM Tris HCl (pH 7.6), 4% [w/v] SDS, 100 mM DTT.

UREA DENATURING BUFFER:

100 mM Tris HCl (pH 8.5), 8M Urea, 100 mM DTT.

nanoHPLC BUFFER A:

0.1% [v/v] formic acid in HPLC-grade H_2O (VWR).

nanoHPLC BUFFER B:

0.1% [v/v] formic acid in HPLC-grade acetonitrile (ACN) (#20060320, VWR).

SCX BUFFER A:

1 mM NH_4HCO_2, 20% [v/v] ACN, 0.05% [v/v] TFA (pH 2.7).

SCX BUFFER B:
200 mM NH_4HCO_2, 20% [v/v] ACN, 1.5% [v/v] TFA (pH 2.7).

GEL FILTRATION BUFFER:
40 mM Tris HCl (pH 7.4), 50 mM NaCl, 5 mM $MgCl_2$.

SAMPLE LOADING SOLUTION:
0.1% [v/v] trifluoroacetic acid (TFA).

2.3 MOLECULAR BIOLOGY

2.3.1 REAGENTS

The following items were from Invitrogen (Paisley, UK), SOC media (#15544-034), ultrapure agarose (#15510-019) and TOPO cloning kit (#450030), as well as subcloning efficiency DH5α (#18265-017), One Shot Mach1 (#C862003) and One Shot BL21 (DE3) (#C600003) chemically competent cells. Luria-Bertani (LB) broth and agar (LAB-168 and LAB-191, respectively) were purchased from LAB M Ltd. (Bury, UK). 100 mM dNTP mix was acquired from Agilent (#200415, Santa Clara, US) and nuclease free water from Sigma (#W4502). The DNA polymerase enzymes utilised were Taq polymerase (#BIO-21040, Bioline, London, UK) and KOD hot start (#71086, Merck Millipore, Billerica, MA, USA). From New England Biolabs (Hertfordshire, UK) were all restriction endonucleases, and Quick load 1kb and 100 bp ladders (#NO648 and #NO467, respectively). National diagnostics provided 50× TAE buffer (#B9-0030). Gel extraction (#28706), Miniprep (#27106) and HiSpeed Maxiprep (#12663) kits were all from Qiagen (Crawley, UK).

2.3.2 PLASMIDS

pCR4-TOPO plasmid was from Invitrogen. pRSET-B-His-N-Ras and pRSET-B-K(B)-Ras (human sequences) were kind gifts from Ignacio Rubio (Institute for Molecular Cell Biology, Friedrich-Schiller-Universität Jena, Germany). Wild-type, full-length human H- and K(A)-Ras sequences were subcloned into the pTrcHis A vector (Invitrogen) from plasmids described in (Laude and Prior, 2008; Prior et al., 2001), as detailed below. All DNA sequences were verified by DNA sequencing services (Dundee, UK). Maps of generated plasmids are provided in supplementary material.

PRIMER NAME	ISOFORM	SEQUENCE	Tm (ºC)
H-Ras_wt_BamH1_F	H-Ras	5'-GAGGATCCATGACGGAATATAAGCTGGTGGTG	64
H-Ras_wt_Xho1_R	H-Ras	5'-GTCTCGAGTCAGGAGAGCACACACTTGCAG	66
K(A)-Ras_wt_BamH1_F	K(A)-Ras	5'-GAGGATCCATGACTGAATATAAACTTGTGGTAG	61
K(A)-Ras_wt_Xho1_R	K(A)-Ras	5'-GGCTCGAGTTACATTATAATGCATTTTTTAATTTTCACAC	60

Table 2.1. Primers used in this thesis. Primers utilised to amplify wild-type H- and K(A)-Ras sequences, from those present in GFP-tagging vectors used in (Laude and Prior, 2008; Prior et al., 2001). Tm indicates melting temperature. F and R denote forward and reverse primers, respectively. Primers were custom ordered from Eurofins MWG operon.

KOD hot start			Taq		
Time (mins)	ºC	No. cycles	Time (mins)	ºC	No. cycles
3	95	1	5	95	1
0.5	95		1	95	
0.5	59	30	1	55	35
1	72		1	72	
5	72	1	5	72	1

Table 2.2. PCR thermocycle programmes utilised in this thesis. Thermocycles were tailored to each polymerase used. KOD was utilised for amplification of Ras sequences for later subcloning, while Taq was employed in the bacterial colony PCR. A 3000 Techne thermocycler was used to perform each PCR reaction.

2.3.3 POLYMERASE CHAIN REACTION (PCR)

H- and K(A)-Ras wild-type human sequences from GFP-tagging vectors were targeted for subcloning into the pTrcHis A his-tagging vector. Primers described in Table 2.1 indicate those used to amplify these Ras sequences and designed to include BamH1/ Xho1 restriction sites. Primers were custom ordered from Eurofins MWG operon (Ebersberg, Germany). Typical PCR reaction volumes were 50 μL; reactions with KOD polymerase were performed in nuclease free water containing 1× enzyme buffer, 0.8 mM dNTPs, 0.2 pM of forward and reverse primers, 1.5 mM $MgSO_4$, 0.5 ng of DNA template (substituted with water in negative PCR control). Mixtures were made up in 0.65 mL Thermowell PCR tubes (#3745, Corning) and subject to thermal cycle programmes detailed in Table 2.2 using a Techne 3000 (Bibby Scientific, Staffordshire, UK). Following the reaction, entire mixtures were subject to agarose gel electrophoresis.

2.3.4 AGAROSE GEL ELECTROPHORESIS

Agarose gels were prepared by mixing ultrapure agarose with 1× TAE buffer (40mM Tris acetate, 1mM Na_2EDTA) before heating in a microwave. Gels containing between 0.8-1% agarose were used in the present thesis. Ethidium bromide was added to a final concentration of 0.5 μg/mL to the gel, which was poured into mould and allowed to set. DNA samples were constituted in 1× DNA loading buffer to volumes of around 15 μL, before being separated in a horizontal midi tank (Invitrogen) at 100-150V for between 1-1.5 hrs. Bands were visualised using a UV dock and excised with a scalpel. DNA was extracted using a Qiagen gel extraction kit, as per manufacturer's instructions.

2.3.5 SUBCLONING AND BACTERIAL TRANSFORMATION

PCR products from above steps were ligated into the pCR4-TOPO vector by mixing 1.5 μL of PCR product (~50 ng), with 1μL salt solution (provided in the TOPO kit), 2.5 μL water and 1 μL of the TOPO vector. The mixture was left for 15 mins at RT, before Mach1 cells were transformed with the ligated vector. The basic transformation procedure of each bacterial strain was as follows: 30 μL of cells were thawed on ice, before <100 ng of plasmid was added and mixed gently. The mixture was incubated on ice for 30 mins, then heat shocked at 42°C for 30 seconds and immediately placed back on ice for 2 minutes. Bacteria were then placed into 250 μL SOC media and allowed to recover at 37°C for 1 hr with shaking at 245 rpm. The mixture was plated out, in varying amounts, on LB agar plates supplemented with kanamycin or ampicillin as appropriate and incubated overnight at 37°C. Bacterial colony PCR was performed on 20 colonies of kanamycin-resistant Mach1 cells for each TOPO vector. Colonies were picked with a toothpick and dipped into the following solution: 1 μL 10× buffer (NH_4), 0.4 μL $MgCl_2$ solution (provided in the Taq polymerase kit), 0.8 mM dNTPs, 0.2 pM of forward and reverse primers, 0.05 μL Taq polymerase, making a final volume of 10 μL. The reaction mixture was subject to the thermocycle protocol stipulated in Table 2.2. This PCR step highlighted colonies containing the desired sequences. Three colonies of Mach1 cells transformed positive for the TOPO vector containing H- and K(A)-Ras sequences were picked again and grown up in LB broth overnight at 37°C with shaking at 245 rpm. DNA from the ciltures was purified using a Qiagen Miniprep kit, as per

manufacturer's instructions. Test digests were performed to further ensure insertion of each sequence into the TOPO vector. 3.5 μL of isolated DNA was mixed with 200U (0.8 μL) of each BamH1 (#r0136, NEB) and Xho1 (#r0146, NEB) restriction enzymes, 2 μL of 10× BSA (provided by NEB), 2 μL of NEB buffer 3 and 10.9 μL water, making a final volume of 20 μL. Digests were performed at 37°C for 2 hours, before being heated at 65°C for 10 minutes to end the reaction. DH5α cells were transformed with the pCR4-TOPO vectors containing H- and K(A)-Ras sequences, as described previously, and a colony of each was grown in 50 mL LB broth supplemented with kanamycin overnight. The culture was subject to Maxiprep DNA purification and the isolated DNA was sequence verified. To subclone the Ras sequences into the pTrcHis A vector, pCR4-TOPO-H-Ras and -K(A)-Ras, as well as the pTrcHis A vector, were digested as described above. The Ras sequences and cut pTrcHis A vector were gel extracted and ligated using T4 ligase (#M0202, NEB), as per manufacturer's instructions. DH5α cells were transformed with the ligated vectors and grown in 50 mL LB broth, as described above, and subject to Maxiprep purification. Insertion of the Ras sequences into the pTrcHis A vector was checked with test digestion, as detailed above.

2.4 RECOMBINANT RAS PROTEIN PRODUCTION

2.4.1 ISOTOPE-LABELLED H-, K(A)-, K(B)- AND N-RAS PROTEIN PRODUCTION

Auxotrophic *E.coli* AT713 bacteria were procured from the Yale Coli Genetic Stock Center (New Haven, USA). This strain lacks ArgA and LysA biosynthetic genes (*argA21, lysA22*), making it suitable to generate isotope-labelled, full-length proteins. To make these bacteria chemically competent, they were grown on non-selective agar overnight at 37°C, before a single colony was picked and incubated in LB broth overnight at 37°C with shaking at 245 rpm. The culture was placed into 100 mL of fresh LB broth and incubated until an optical density (OD) of 0.4 was reached, as measured using a Multiskan Spectrum (Thermo Scientific, Rockford, IL, USA). Bacteria were pelleted by centrifugation at 2, 500 rpm for 10 min at 4°C, supernatant discarded and pellet resuspended in 50 mL of the following ice-cold, sterile filtered media: 80 mL LB broth, 10% [w/v] PEG, 10 mM $MgCl_2$, 10 mM $MgSO_4$ and 5 mL DMSO. Cells were chilled on ice for 10 mins before being aliquoted and stored at -80°C. Chemical competency was checked by transforming the cells with pRSET-B-K(B)-Ras and growing cells on LB agar supplemented with ampicillin. The

transformation procedure of the AT713 cells was the same as that stipulated in section 2.3.5, except 50 μL of cells were used.

AT713 bacteria were transformed with pRSET-B-K(B)-RAS, pRSET-B-N-RAS, pTrcHis-A-H-Ras and pTrcHis-A-K(A)-Ras plasmids and grown in 5 mL LB broth for 6 hours at 37°C with shaking at 245 rpm. 100 μL this culture was then transferred to 100 mL of M9 minimal media, supplemented with L-lysine-U-$^{13}C_6$-$^{15}N_2$ (Lys8) and L-arginine-U-$^{13}C_6$-$^{15}N_4$ (Arg10) both from Sigma (#608041 and #608033, respectively), and grown overnight under the same conditions as in the LB broth. 10 mL of culture was then spiked into 10× 100 mL fresh M9 minimal media, with heavy isotopes of arginine and lysine, and allowed to grow as before. When an OD of 0.6 was reached, isopropyl β-D-1-thiogalactopyranoside (IPTG, #I6758, Sigma) was added to a final concentration of 1 mM and the cells incubated for a further 3 hours. Bacteria were collected and pelleted by centrifugation at 4, 500 rpm, washed in ice-cold PBS and pelleted once again.

2.4.2 RAS PROTEIN PURIFICATION

His-tagged Ras isoforms were subject to purification using a PrepEase his-tag purification kit (#78791 1 KT, USB now part of Affymetrix, Santa Clara, USA). Fresh bacterial AT713 cell pellets were resuspended in ice-cold 1× LEW buffer (supplied in kit), supplemented with bacterial protease inhibitors (diluted 1:500, #P8465, Sigma-Aldrich). The mixture was then frozen at -80°C overnight and allowed to thaw on ice the following morning, before lysozyme (#L6876, Sigma-Aldrich) was added to a final concentration of 1 mg/ mL and incubated on ice for 30 mins with regular agitation. The culture was then sonicated 10 times for 20 seconds with 30-second intervals on ice, ahead of centrifugation at 40, 000 rpm for 30 mins at 4°C. The subsequent supernatant was subject to the remainder of the PrepEase his-tag purification kit, as per manufacturer's instructions. Eluted Ras protein was concentrated using an Amicon Ultra 15-kDa MWCO filter (#UFC900, Millipore) through centrifugation at 4, 000 rpm until the total volume was reduced to ~1 mL. Ras elutions were then subject to gel filtration using an ÄKTA purifier and Superdex 200 10/300 GL, both from GE Healthcare (Buckinghamshire, UK). Gel filtration buffer, stipulated in section 2.1, was passed through the column at 0.5 mL/ min and 0.5 mL fractions were collected each minute. The purity of Ras isoforms was checked by subjecting ~1 μg of each protein to SDS-PAGE, using 4-12% [w/v] bis-

tris NuPAGE gels (#NP0335BOX, Invitrogen), before colloidal blue staining (#LC6025, Invitrogen) and visualisation on an Odyssey system (LI-COR, NB, USA). Ras PSAQ standards were flash frozen in liquid nitrogen and stored at -80°C in LoBind tubes (#022431081, Eppendorf, Hamburg, Germany) that were pre-treated with 2% [w/v] BSA for 2 hours, before all BSA was removed and excess liquid allowed to evaporate in a tissue culture hood for 3 hours. All subsequent dilutions of stock protein standard was performed in gel filtration buffer supplemented with 2% [w/v] BSA.

2.4.3 DETERMINATION OF ISOTOPE INCORPORATION IN THE RAS PSAQ STANDARDS

500 ng of each isotope-labelled Ras protein was subject to SDS-PAGE and in-gel digestion with trypsin or elastase, before SRM-analysis on a 4000 QTRAP (AB SCIEX, Warrington, UK). The total wild-type Ras peptide, LVVVGAGGVGK, and pan-Ras peptide, FYTLVREI, were monitored in both light and heavy forms. The differences in signals between light and heavy was used as readout on the extent of isotope labelling. The materials and procedures involved in this analysis are described in depth in section 2.6.

2.4.4 DETERMINATION OF ISOTOPE-LABELLED RAS PROTEIN CONCENTRATION

Accurate measurement of the isotope-labelled, his-tagged Ras proteins is essential for the success of the PSAQ analysis. Isotopically light, his-tagged K(B)-Ras was generated from AT713 bacteria, as detailed above, but using LB broth instead of M9 minimal media. The K(B)-Ras protein was purified with the PrepEase kit and gel filtrated using phosphate-buffered saline (PBS, #70011044, Invitrogen). The light K(B)-Ras protein was subject to a bicinchoninic acid (BCA) assay (#23225, Thermo), using the supplied BSA standards (#23208, Thermo), which, according to the manufacturer, is precisely formulated at 2 mg/mL and calibrated by direct comparison to purified BSA (Fraction V) from the National Institute of Standards and Technology (NIST). A second independent protein concentration measurement was made using an absorbance measurement at 280 nm on a Nanodrop 1000 (Thermo). PBS will not interfere with absorbance readings at this wavelength, making it a suitable buffer for such measurements. The molar extinction coefficient for the K(B)-Ras protein was calculated as 13660 M^{-1} cm^{-1} using ProtParam (http://web.

expasy.org/protparam/) and verified manually using the following equation (Gill and von Hippel, 1989; Pace et al., 1995):

$$\varepsilon\,(280)(M^{-1}cm^{-1}) = (\#Trp)(5{,}500) + (\#Tyr)(1{,}490) + (\#cystine)(125)$$

Isotope-labelled Ras standards had their concentrations measured using a Bradford assay (Bio-Rad, Hertfordshire, UK), before they were mixed with 500 or 250 ng of the accurately quantified light K(B)-Ras at ratios between 1-2.5. The protein mixes were then subject to SDS-PAGE, in-gel digestion and subsequent SRM-analysis, detailed in section. The total wild-type Ras peptide, LVVVGAGGVGK, was monitored for both light and heavy versions enabling the accurate and precise quantification of the isotope-labelled Ras PSAQ standards. Detailed information on these procedures is provided in section 2.6.

2.5 CELL LINES, GENERAL CELL CULTURE AND PROTEIN BIOCHEMISTRY

2.5.1 CELL LINES, GENERAL CULTURE AND CELL LYSIS

HeLa cells were from Graham Warren (Max F. Perutz Laboratories, Vienna, Austria) and grown in Dulbeccco's Modified Eagles' Medium (DMEM) with GlutaMAX (#10566, Invitrogen), supplemented with 10% [v/v] foetal bovine serum (FBS, #16000, Invitrogen), 100 U/ mL penicillin and 100 μg/ mL streptomycin (#15240, Invitrogen). Isogenic SW48 cell lines were from Horizon Discovery Group plc (Cambridge, UK): Parental (#HD PAR-006), heterozygous knock-in cells containing $KRAS^{G12D}$ (#HD 103-011), $KRAS^{G12V}$ (#HD 103-007), $KRAS^{G13D}$ (#HD 103-002), $KRAS^{G12C}$ (#HD 103-006), $KRAS^{G12S}$ (#HD 103-013), $KRAS^{G12A}$ (#HD 103-009), $KRAS^{G12R}$ (#HD 103-010), $HRAS^{12V}$ (#HD 103-034), $NRAS^{G12V}$ was generated in house by Dr. Simon Oliver and Prof. Ian Prior (University of Liverpool). Second, independent SW48 isogenic clones: $KRAS^{G12D}$, $KRAS^{G12V}$, $KRAS^{G13D}$ share the same catalogue numbers as the first clones. All heterozygous knock-in cell lines were genetically matched to the Parental cell line. SW48 cells were grown in McCoy's 5A medium with GlutaMAX (#36600, Invitrogen), supplemented with 10% [v/v] FBS and penicillin/ streptomycin as for DMEM. All cells were grown at 37°C with 5% [v/v] CO_2 in a Binder cell incubator (Tuttlingen, Germany) and, unless stipulated otherwise, in 100 mm cell culture dishes (#430167, Corning). A confluent

monolayer (80% or higher) in a 100 mm dish was washed twice with 5 mL of ice-cold PBS, before 200 μL of lysis buffer (RIPA or NP40) was added. Cells were rocked on ice for 15 minutes, before lysates were collected and cleared by centrifugation at 13, 000 rpm for 15 min at 4°C. Lysates were subject to a BCA assay (Thermo), as per manufacturer's instructions. 100, 50 or 25 μg of lysate was constituted in 1× sample buffer, heated for 5 minutes at 95°C, allowed to cool and centrifuged briefly, before being subject to SDS-PAGE. RIPA was used for Ras abundance analysis, while NP40 lysis buffer was used for the remaining immunoblot experiments.

2.5.2 SDS-PAGE AND IMMUNOBLOTTING

Protein gels were either bought from Invitrogen (NuPAGE range) or were hand-made using a Bio-RAD system with ingredients supplied by National Diagnostics: 4× ProtoGel resolving buffer (#EC-892), ProtoGel stacking buffer (#EC-893), ProtoGel (40% [v/v]) (#EC-891), N,N,N',N'-tetramethylethylene diamine (#EC-503). Ammonium persulphate was from Sigma (#A3678). Gels containing 10% [v/v] acrylamide were used to separate most lysates, while those containing 15% [v/v] acrylamide used for immunoblot analysis of cellular Ras abundance. Gels were run using Bio-RAD or NuPAGE-specific gel tanks at 50V for the first 15 mins, then 150V thereafter. Molecular weight markers were rainbow marker (#RPN800E, GE Healthcare) and pre-stained protein marker, broad range (#P7706, NEB). Proteins were transferred from gels to 0.22 or 0.45 μm nitrocellulose membranes using a GENIE blotter full submersion apparatus (Idea Scientific, MN, USA) at 0.9 amps for 1 hr. For immunoblot analysis of Ras proteins, 0.22 μm pore size nitrocellulose membranes were used, whereas 0.45 μm was used for all other proteins. Following transfer, membranes were stained with Ponceau S solution (#P1710, Sigma) to visualise protein content and removed with washing in PBS. Membranes were blocked with Marvel milk, fish skin gelatin or BSA, before addition of primary antibody in blocking solution as appropriate (primary antibodies employed in this thesis are detailed in Table 2.3). Membranes were washed 3 times for 5 minutes using TBS-T, ahead of addition of infrared fluorescent secondary antibodies (LI-COR) for 1 hr in blocking solution in the dark. Membranes were washed again, as before, then analysed using an Odyssey system (LI-COR).

SPECIFICITY	CAT NO.	COMPANY	SPECIES	BLOCK	DILUTION
ERK	4695	CST	Rabbit	BSA	1:1,000
pERK (Thr202/Tyr204)	4370	CST	Rabbit	BSA	1:500
MEK	9122	CST	Mouse	BSA	1:500
pMEK (ser217/221)	9154	CST	Mouse	BSA	1:500
AKT	2920	CST	Mouse	BSA	1:500
pAKT (ser473)	9271	CST	Mouse	BSA	1:500
B-RAF	9434	CST	Mouse	Marvel	1:1,000
C-RAF	9422	CST	Rabbit	Marvel	1:1,000
EGFR	sc-03	SC	Sheep	Marvel	1:500
MET	3127	CST	Mouse	Marvel	1:400
H-Ras	sc-520	SC	Rabbit	Marvel	1:200
K-Ras	R300	Sigma	Mouse	Gelatin	1:100
N-Ras	sc-31	SC	Mouse	Marvel	1:200
Pan-Ras	Y13259	Sigma	Rat	Gelatin	1:100
ZO-2	2847	CST	Rabbit	Marvel	1:1,000
AKAP12	GTX106373	GeneTex	Rat	Marvel	1:2,000
DCLK1	SAB4200186	Sigma	Rabbit	Marvel	1:1,000
Caveolin-1	610060	BD	Rabbit	Marvel	1:1,000
Tubulin	Ab6276	Abcam	Mouse	Marvel/ BSA	1:10,000
Actin	T5168	Bethyl	Mouse	Marvel/ BSA	1:10,000

Table 2.3. Primary antibodies employed in the present thesis. CST, Cell Signalling Technology, Leiden, The Netherlands; SC, Santa Cruz Biotechnology, Heidelberg, Germany; GeneTex, Hsinchu City, Taiwan; Sigma, Sigma-Aldrich; BD, BD Transduction Laboratories, Oxford, UK; Bethyl, Bethyl Laboratories Inc., TX, USA.

Secondary antibodies employed were the IRDye range from LI-COR, specifically: donkey anti-mouse IRDye 800cw (#926-32212), donkey anti-mouse IRDye 680cw (#926-32222), donkey anti-rabbit IRDye 800cw (#926-32213), donkey anti-rabbit IRDye 680cw (#926-32223), donkey anti-sheep IRDye 800cw (#926-32214), donkey anti-rat IRDye 800cw (#926-32219). All were diluted 1:15, 000 in blocking buffer (BSA or Marvel) before use.

2.6 SRM ANALYSIS OF CELLULAR RAS ABUNDANCE

2.6.1 DETECTION OF PROTEOTYPIC RAS PEPTIDES

1 μg of isotope-labelled Ras protein isoforms were constituted in 1× sample buffer, before being subject to SDS-PAGE using NuPAGE NOVEX 10-well, 1.5 mm 4-12% bis-tris gels (#NP0335BOX, Invitrogen). Running of the gel was as described earlier. Gels were fixed with 50 mL of 10% acetic acid (#A6283, Sigma), 40% HPLC water, 50% HPLC methanol (#20864, VWR) for 15 mins at RT, before staining with colloidal blue and excision of his-tagged Ras proteins with a scalpel. Gel pieces were diced into ~1 mm^3 cubes, placed into LoBind tubes and destained using 50 mM ambic, 50% ACN with heating and shaking at 37°C and 900 rpm using a Thermoshaker (Thermo). Once destained, gel pieces were reduced with reduction buffer for 30 mins at 37°C and 900 rpm, before the gel pieces were allowed to cool to RT and alkylated with alkylation buffer performed at RT in the dark for 30 mins. Gel pieces were then washed by dehydrating with 100% ACN and rehydrating with HPLC water two times. Gel pieces were dehydrated, ahead of addition of trypsin Gold (#V5280, Promega, Madison, WI, USA), LysC (#125-02543, Wako Chemicals, Richmond, VA, USA) or Elastase (#342682, Calbiochem, now part of Merck Millipore). Trypsin and LysC were used at a 1:25 protein-to-protein concentration (40 ng), while elastase was used at a 1:133 ratio (7 ng). All proteases shared the same reaction buffer: 40 mM ambic, 9% [v/v] ACN. Each digest was performed overnight at 37°C. The following morning, supernatants were removed, placed into LoBind tubes and formic acid added to a final concentration of 0.1% to end the digestion. Gel pieces were dehydrated with ACN for 15 mins at RT with shaking at 900 rpm, then the supernatant added to the previous one. Peptides were spun down at 15, 000 rpm for 10 mins, then transferred to fresh LoBind tubes, leaving behind 5 μL that may contain gel debris. Peptides were then dried for ~ 3 hrs in a vacuum concentrator at 60°C (RVC 2-25, Christ, Osterode am Harz, Germany). Peptides were resuspended in 0.1% TFA, placed into sample vials (#186000385c, Waters, Hertfordshire, UK) and subject to LC-MS/MS analysis.

~100 ng of peptides were separated using a nanoAcquity UPLC system (Waters), with a C18, 1.7 μm particle size, 75 μm × 250 mm BEH130 column (#186003545, Waters) and C18, 5 μm particle size, 180 μm × 20 mm, Symmetry, 2G-V/M trap column (#186006527, Waters). Buffers used described in section 2.2. Peptides were

trapped for 1 min at 15 μL/ min flow rate in 0.1% [v/v] formic acid and 0.1% [v/v] ACN, before being separated on a 20 min, 3 to 62.5% CAN [v/v] linear gradient at a flow rate of 0.4 μL/ min. The column was kept at a constant 65°C. Eluting peptides were ionised by a nano-electrospray ionisation source (Proxeon, now part of Thermo) before MS analysis on an LTQ Orbitrap XL in data-dependent acquisition mode (DDA) (Thermo Fisher Scientific, Bremen, Germany). The Orbitrap acquired MS survey scans at 30, 000 resolution and *m/z* range of 300 to 2000. MS/MS of the top 5 multiply charged ions were acquired in the LTQ, after isolation and CID for 30 ms at 35% energy. Dynamic exclusion was imposed with a rolling exclusion list for 180 secs (n = 1). All spectra were acquired using Xcalibur software (version 2.0.7; Thermo Fisher Scientific). Acquired .RAW files were converted to the .mgf file format using msconvert (http://proteowizard.sourceforge.net/tools/msconvert.html) then subject to a Mascot MS/MS ion search (Matrix Sciences, London, UK) using the SwissProt human database with carbamidomethylation of cysteine set as a fixed modification, with 1 missed cleavage and peptide charge states of +2, +3 and +4. For elastase, no enzyme was selected. MS utilising the LTQ Orbitrap was performed at the University of Liverpool.

~100 ng of peptides were separated using a nanoAcquity, as described above, and ionised using a nanospray II source (AB SCIEX) before MS analysis on a 4000 QTRAP in information-dependent mode (IDA). An enhanced mass spectrometry (EMS) scan at an *m/z* range of 400 to 1200 at 4, 000 Da/ sec scan rate was performed, before an enhanced resolution (ER) scan at 200 Da/ sec of the top 5 multiply charged ions that exceed 100, 000 counts per second (cps). The top 2 ions were then fragmented in the collision cell using CID with a rolling collision energy (CE), calculated using the equation:

$$CE\ (eV) = (slope \times m/z) + intercept$$

where slope and intercept are 0.058 and 4, respectively, for singularly charged ions, 0.044 and 3 for +2 ions, and 0.044 and 2 for +3 ions. MS/MS spectra were obtained using an enhanced product ion (EPI) scan at 4000 Da/ sec using an *m/z* range of 100 to 1100 Th. Dynamic fill time was employed to determine accumulation time for fragment ions. All spectra were acquired using Analyst (version 1.5, AB SCIEX). Acquired .wiff files were converted into the .mgf file format and searched as above.

MRM-initiated detection and sequencing (MIDAS) (Unwin et al., 2005) was performed with an initial user-defined MRM stage, before ions detected over a 10, 000 cps intensity threshold were fragmented, as described above, followed by an EPI scan. Dynamic fill time was also utilised. Acquired MS/MS spectra were inspected manually. MS using the 4000 QTRAP was performed at the Paterson Institute for Cancer Research (now The Manchester Institute), University of Manchester, UK.

2.6.2 SELECTION OF PROTEOTYPIC RAS PEPTIDES

Ras proteotypic peptides were selected according to the criteria specified in section 3.1.2.

2.6.3 SELECTION OF TRANSITIONS

Transitions were selected following the criteria set out in section 3.1.3. Table 3.1 describes all transitions and 4000 QTRAP instrument settings utilised in the present thesis.

2.6.4 COLLISION ENERGY OPTIMISATION

Each peptide had a CE calculated using the equation described in section 2.6.1. ~100 ng of peptides were separated using a nanoAcquity system using a gradient described in the next section. Peptides were then analysed using SRM on a 4000 QTRAP with collision energies -4, -2, 0, +2 and +4 eV from the calculated CE value. Q3 values were adjusted by 0.1 Da for each different CE tested, as in (Sherwood et al., 2009). Optimal CE values were those that generated the greatest peak heights in 3 independent replicates.

2.6.5 RETENTION TIME DETERMINATION

IDA of 200 ng of Mass PREP *E.coli* digestion standard (#186003196, Waters) was performed to determine at what volume percentage of ACN peptides first and last eluted. A variety of different gradients were then used to separate ~100 ng of Ras peptides with SRM performed using calculated, non-optimised CE for each peptide.

A 60-min 8-65% ACN in 0.1% [v/v] formic acid linear gradient was selected, following the criteria stipulated in section 3.1.4. Only differences from section 2.6.1 were that trapping flow rate was 3 μL/ min for 10 mins and column temperature was 60°C. Retention times were frequently assessed, since they change as the column and trap mature. The retention times detailed in Table 3.2 represent Ras peptide retention times on a fresh column and trap.

2.6.6 VALIDATION OF INSTRUMENT PARAMETERS

All optimised instrument settings were tested to ensure each Ras peptide generated a linear response between peptide abundance and measured intensity on the 4000 QTRAP. Light H-, K(B)- and N-Ras proteins were produced in One Shot BL21 cells, following the protocols described above, gel filtrated in PBS and their concentration measured using a BCA assay. These proteins were also utilised in the immunoblot analysis of cellular Ras abundance. 250 ng of light Ras isoform protein was mixed with its counterpart isotope-labelled Ras PSAQ standard at ratios from 0.2 to 20. All dilutions were performed in gel filtration buffer supplemented with 2% [w/v] BSA. The protein mixtures were constituted in 1× sample buffer, before being subject to in-gel digest as in section 2.6.1. Dried peptides were resuspended in 25 μL 0.1% [v/v] TFA and 5 μL of sample was subject to SRM analysis as described in section 2.6.10.

2.6.7 CELL COUNTING AND LYSIS

The SW48 cell lines, described in section 2.5.1, were grown to ~80% confluency before being washed with warm PBS, trypsinised and resuspended in 10 mL of PBS. 10 μL of cells was placed on either side of a Bright-Line haemocytometer (#Z359692, Sigma-Aldrich) and cells within 3 squares on either side were counted. 9 mL of cells were spun down at 200 rcf for 5 min at RT, supernatant was checked to ensure no clumps of cells were floating and the PBS was carefully aspirated. 200 μL of ice-cold RIPA buffer (without PhosSTOP tablets and NaF) was added to the cell pellet and broken apart by briefly pipetting up and down. Cells were then incubated on ice for 20 mins, with regular slight agitation. Lysate volume was measured using a 100 μL Hamilton syringe (Hamilton Bonaduz AG, Bonaduz, Switzerland). Cell counts and volume measurements are provided in the supplementary material,

which enable the number of cells used in the PSAQ assay to be determined. Lysates were spun at 13, 000 rpm for 15 mins at 4°C. Supernatant was taken, diluted 1:10 in RIPA buffer and subject to a BCA protein assay. Cleared lysates were flash frozen in liquid nitrogen and stored at -80°C until needed.

2.6.8 SDS-PAGE AND IN-GEL DIGESTION OF PSAQ TARGET LYSATES

100 μg of cell lysates were constituted in 1× sample buffer, along with isotope-labelled PSAQ Ras standards (10 ng of K(B)-Ras, 5 ng of H- and N-Ras), before being subject to SDS-PAGE as described in section 2.6.1. Rainbow marker (GE Healthcare) was used as a protein marker. The gel was only fixed, not stained, to avoid excess colloidal blue blocking chromatography columns. A gel band spanning 31-17 kDa was excised and diced, followed by in-gel digestion, as described in section 2.6.1, but using 400 ng of trypsin.

2.6.9 SELECTED REACTION MONITORING AND DATA ANALYSIS

~1 μg of peptides were separated using a nanoAcquity system using optimised CE values and chromatography gradients, as described in the previous sections. At the start of each run, an IDA scan was performed for the first and last 10 mins to obtain a basic assessment of spray performance. Between every 5 samples, or if a problem was observed in samples, an IDA scan of 100 ng of *E.coli* digest standard (Waters) was performed to determine chromatography performance. The acquired *E.coli* data were checked for consistent intensities and numbers of peptides identified in a Mascot MS/MS ion scan. If a fall in 40% intensity or 20% identified peptides was observed, the column and trap were exchanged, along with cleaning of the ion source. As mentioned in section 2.6.6, the retention times of each peptide will change with usage. In Table 2.4, example acquisition windows for each peptide are provided. These windows were altered as necessary to accommodate for shifts in retention time. Two independent acquisitions were required per sample, since H-Ras specific, SYGIPYIETSAK, and N-Ras specific, SFADINLYR, eluted at very similar times. Data analysis was performed using Skyline (version 2.4) (MacLean et al., 2010).

RUN 1			RUN 2		
ISOFORM	PEPTIDE	WINDOW (mins)	ISOFORM	PEPTIDE	WINDOW (mins)
-	IDA	0 - 10	-	IDA	0 - 10
total wt	LVVVGAGGVGK	11 - 36	total wt	LVVVGAGGVGK	11 - 34
H	SYGIPYIETSAK	37 - 45	N	SFADINLYR	35 - 44
K(B)	QGVDDAFYTLVR	45 - 55	K + N	SYGIPFIETSAK	45 - 55
H + K	TGEGFLCFAINNTK	55-75	H + N	QGVEDAFYTLVR	55 - 72
-	IDA	remainder	-	IDA	remainder

Table 2.4. Acquisition time windows for proteotypic Ras peptides during the SRM-based quantitative analysis. IDA was performed before and after SRM analysis to obtain a basic assessment of spray performance. A 60-min, 8-35% ACN linear gradient was used to separate Ras peptides, with total liquid chromatography run time at 105 mins, including regeneration and re-equilibration steps. IDA, information dependent acquisition.

MUTATION	PEPTIDE	CHARGE	PRECURSOR (*m/z*)	FRAGMENT (*m/z*)	ION
G12D	LVVVGA**D**GVGK	+2	507.3	801.4	y9
				702.4	y8
				603.3	y7
G12V	LVVVGA**V**GVGK	+2	499.3	785.5	y9
				686.4	y8
				587.4	y7
G13D	LVVVGAG**D**VGK	+2	507.3	801.4	y9
				702.4	y8
				603.3	y7
G12C	LVVVGA**C**GVGK	+2	529.8	846.4	y9
				747.4	y8
				648.3	y7
G12S	LVVVGA**S**GVGK	+2	493.3	773.5	y9
				674.3	y8
				575.3	y7
G12A	LVVVGA**A**GVGK	+2	485.3	757.5	y9
				658.4	y8
				559.3	y7
G12R	LVVVGA**R**GVGK	+2	527.8	842.5	y9
				743.5	y8
				644.4	y7

Table 2.5. Targeted precursor and fragment ion masses of Ras mutant-specific peptides. Peptide precursor and fragment ion masses were calculated using Skyline (version 2.4). Sample preparation and MIDAS settings are described in section 2.6.1 and 2.6.11. MIDAS, MRM-initiated detection and sequencing.

To calculate copy numbers per cell, ng values of Ras abundance were converted into moles, before being divided by the number of cells that produced the 100 μg of lysate subject to SRM analysis. The subsequent molar amount was then multiplied by the Avogadro constant (6.0221413×10^{23}). Results were visualised using JMP10 (SAS, Cary, NC, USA).

2.6.10 MIDAS DETECTION OF MUTANT-SPECIFIC RAS PEPTIDES

Precursor and fragment ion masses of mutant Ras peptides, detailed in Table 2.5, were calculated using Skyline (version 2.4). Isogenic SW48 cells, harbouring the various K-Ras mutant proteins, were subject to sample preparation as described earlier. All lysates were subject to in-gel digestion with trypsin, except for the K-RasG12R cell line where LysC was used. To ensure selectivity, only y ions were targeted in the MIDAS analysis. Targeted ions are described in Table 2.5. The MIDAS settings were as those described in section 2.6.1.

2.6.11 DYNAMIC SILAC ANALYSIS OF RAS PROTEIN TURNOVER

Parental SW48 cells, and those harbouring heterozygous $KRAS^{G12D}$, $KRAS^{G12V}$ and $KRAS^{G13D}$ mutations, were grown for 7 passages in SILAC DMEM (#LM010, Dundee cell products, Dundee, UK), containing either 'light' lysine (Lys0) and arginine (Arg0) (#L5501 and #A5506, respectively, Sigma) or 'medium' L-lysine-$^{2}H_4$ (Lys4) and L-arginine-U-$^{13}C_6$ (Arg6) (#616192 and #643440, respectively, Sigma). All SILAC media also contained 200 mg/ ml proline (#P0380, Sigma) and 10% [v/v] dialysed FBS. Full FBS (Invitrogen) was dialysed using SnakeSkin 10 kDa MWCO dialysis tubing (#88245, Thermo) in 5 L of PBS over 24 hrs × 2 at 4°C, before sterile filtering (#431097, Corning). Labelling of cells was checked by subjecting 50 µg of lysate from each labelled cell line to in-gel digestion with trypsin, before SRM analysis on a 4000 QTRAP monitoring for the total wild-type Ras peptide, LVVVGAGGVGK, and the K(B)-Ras specific peptide, QGVDDAFYTLVR.

4 million cells were plated into 100 mm dishes and allowed to attach for 12 hrs. 'Medium' cells were then washed with PBS and exposed to 'heavy' SILAC DMEM containing L-lysine-U-$^{13}C_6$-$^{15}N_2$ (Lys8) and L-arginine-U-$^{13}C_6$-$^{15}N_4$ (Arg10) for 0.5, 4, 7, 11, 27 or 48-hour time points. All cells were trypsinised, spun down at 200 rcf and lysed as in section 2.6.8. 50 μg of 'light' cells was mixes with 50 μg of 'medium/

heavy' cells and subject to SDS-PAGE and in-gel digestion as in section 2.6.9. Utilised transitions are detailed in Table 3.3. Data analysis was performed using Skyline (version 2.4). 50% turnover was determined to be the point where heavy and medium signals reached a ratio of 1. For further details see section 3.2.5.

2.7 PROTEOMIC AND PHOSPHO-PROTEOMIC PROFILING OF ISOGENIC SW48 CELLS

2.7.1 CELL CULTURE, SILAC MEDIA PREPARATION AND CELL LYSIS

Parental SW48 cells, and those harbouring heterozygous $KRAS^{G12D}$, $KRAS^{G12V}$ and $KRAS^{G13D}$ mutations, were grown for 7 passages in SILAC McCoy's 5A media (custom preparation, Dundee cell products), containing either 'light' lysine (Lys0) and arginine (Arg0), 'medium' L-lysine-$^{2}H_4$ (Lys4) and L-arginine-U-$^{13}C_6$ (Arg6) or 'heavy' L-lysine-U-$^{13}C_6$-$^{15}N_2$ (Lys8) and L-arginine-U-$^{13}C_6$-$^{15}N_4$ (Arg10). Each media also contained 200 mg/ mL of proline and 10% [v/v] of self-dialysed 10 kDa MWCO FBS. At least three biological replicate data sets of each $KRAS^{MUTANT}$ versus Parental SW48 were obtained (n = 4 for $KRAS^{G12D}$ versus Parental). Figure 3.26 provides an overview of the experimental setup utilised in this thesis. Incorporation of isotopes was checked by subjecting 25 μg of each cell line lysate to SDS-PAGE and in-gel tryptic digestion, before LC-MS/MS analysis on the LTQ Orbitrap XL in DDA mode, as described above. Acquired .RAW files were analysed using MaxQuant (version 1.0.13.13), followed by quantification of incorporation using an R script provided by Christian Kelstrup (Faculty of Health and Medical Sciences, University of Copenhagen, Denmark).

For SILAC experiments, cells were grown to ~80% confluency in 245 cm^2 square dishes (#431111, Corning), before being washed twice in ice-cold PBS and scraped into 10 mL of PBS per dish. Cells were combined where appropriate and spun down at 200 rcf for 15 min at 4°C. Supernatant was removed and cells were resuspended in ice-cold NP40 buffer, also supplemented with 40 mM β-Glycerol-Phosphate, 1 mM EDTA and 1mM Na_3VO_4, in addition to the supplements described in section 2.1. Cells were incubated on ice for 15 mins before being spun down again as above. The supernatant was taken and used as a crude cytosolic fraction. The pellet was resuspended in SDT buffer and treated as a crude nuclear fraction. Protein content was assessed using a BCA assay. 15 mg of each 'light', 'medium' and

'heavy' cell lysates were then mixed (1:1:1 protein ratio).

2.7.2 SAMPLE PREPARATION FOR PROTEOME ANALYSIS

50 μg of the 45 mg 1:1:1 protein mix was constituted in 1× sample buffer, before being subject to SDS-PAGE and protein visualised through staining with colloidal blue. The gel was cut into 48 slices with a scalpel, according to protein content, and each slice diced, destained and in-gel tryptic digested as described earlier. Dried peptides were resuspended in 25 μL 0.1% [v/v] TFA and 5 μL was subject to LC-MS/MS analysis on a LTQ Orbitrap XL in DDA mode, described further in section 2.7.4.

2.7.3 FILTER-AIDED SAMPLE PREPARATION, STRONG CATION EXCHANGE, TiO_2-BASED PHOSPHOPEPTIDE ENRICHMENT

For phosphopeptide isolation, filter-aided sample preparation (FASP) (Wisniewski et al., 2009) was utilised, before strong cation exchange (SCX) and TiO_2-based phosphopeptide enrichment, based on (Olsen et al., 2006; Olsen and Macek, 2009). The 45 mg protein mix was precipitated by adding 100% TCA at a 1:5 ratio and incubating on ice overnight. The following morning, protein precipitates were pelleted by centrifugation at 1, 000 rpm and the supernatant discarded. Protein pellets were twice washed with ice-cold acetone for 5 mins. Pellets were then resuspended in urea denaturing buffer and sonicated for 10-20 times for 20 sec with 30 sec intervals on ice when necessary. Amicon Ultra 30 kDa, 15 mL filters (#UFC903024, Millipore) were equilibrated by spinning 5 mL HPLC water through the filter at 4, 500 rpm. Resuspended protein was placed into 3 separate filters and spun down until the volume was reduced to 0.5 mL. 10 mL of fresh urea denaturing buffer was added to each filter and spun down again. To the concentrated 0.5 mL of protein, 3 mL of 55 mM IAA in urea denaturing buffer was added, followed by a 30 min incubated at RT in the dark. The filters were then spun as above until the volume was again reduced to 0.5 mL, then a further 10 mL of urea buffer was added and spun once again. 10 mL of 55 mM ambic was added to each filter and spun down to 0.5 mL; this was repeated twice. 2 mL of ambic containing trypsin at a 1:300 protein ratio was added to the filters and placed into a wet chamber (50 mL falcon with moist tissue at the bottom) and incubated overnight at 37°C. The following morning, filters were spun in fresh 50 mL falcon tubes until 0.2 mL was

remaining. 2 mL of ambic was added and spun down again; this was repeated twice. The ambic flow throughs were combined and acidified to a 1% final TFA concentration. Samples were desalted by applying them OASIS HLB cartridge (#186000117, Waters) and peptides eluted with 10 mL 90% ACN, 0.5 % formic acid, before drying in a vacuum concentrator at 60°C. Dried peptides were resuspended in 2 mL of SCX buffer A and subject to strong cation exchange (SCX) using an ÄKTA system with attached S Resource column (#17-1178-01, GE Healthcare). Peptides were bound to the SCX column and eluted with increasing concentrations of SCX buffer B. A three step gradient was utilized at a flow rate of 1 mL/ min: (1) 0% [v/v] B was applied to the column between for the first 17.5 mins following injection of peptides onto the column, (2) a 20 min, linear gradient from 0 to 50% [v/v] B, (3) 50 to 100% [v/v] B in 5 mins (Beausoleil et al., 2004; Hernandez-Valladares et al., 2014; Olsen et al., 2006). 1 mL fractions were collected each minute and dried in a vacuum concentrator as above. Peptides were then subject to TiO_2-based enrichment using stage tips, as described in (Dulla et al., 2010; Rappsilber et al., 2007; Sugiyama et al., 2007). Micro-columns were prepared by placing a small plug of C8-silica matrix (#12145002, Varian, Postfach, Germany) into 200 μL pipette tips (#S1111-1806, Starlab, Muri, Switzerland) and pushing them tight to the end. Peptides were dissolved in 30 μL 80% [v/v] ACN, 2% [v/v] TFA. 500 mg of Titansphere 10 μm TiO_2 beads (#5020-75010, GL Sciences, Eindhoven, The Netherlands) were resuspended in 5 mL of 80% [v/v] ACN, 0.2% [v/v] TFA. 50 μL of bead slurry was added to the micro-columns, followed by two washes of 50 μL 0.6% [v/v] suprapure NH_4OH (#105428, Merck) with centrifugation at 3, 000 rpm for 5 mins. This was repeated with 50 μL glycolic acid (GA) buffer (80 mg/ mL GA in 80% [v/v] ACN, 2% [v/v] TFA). GA was from Sigma (#G8284, Sigma). Peptides were applied to the columns and spun such that it took at least 5 mins for the liquid to pass through. Flow throughs were reapplied to the column and flicked vigorously to mix, before spinning down as before. TiO_2 beads were then washed with 40 μL GA buffer, followed by 40 μL [v/v] 80% ACN, 0.2% [v/v] TFA and 40 μL 20% [v/v] ACN. Phosphopeptides were first eluted into LoBind tubes using 40 μL 0.6% suprapure NH_4OH, followed by 40 μL 60% [v/v] ACN. The second elution involved 40 μL 1% [v/v] pyrrolidine (#83241, Fluka now part of Sigma-Aldrich), followed by 40 μL 60% [v/v] ACN. Both first and second elutions were dried using a vacuum concentrator as described earlier. Phosphopeptides were then resuspended in 0.5% [v/v] formic acid for MS analysis.

2.7.4 MASS SPECTROMETRY ANALYSIS, SPECTRAL MATCHING AND QUANTIFICATION

Peptides were separated using a 60-min, 3-65% ACN in 0.1% formic acid gradient using a nanoAcquity and chromatography columns as described earlier. Proteome samples, described in section 2.7.2, were analysed using an LTQ Orbitrap XL in DDA mode, as described in section 2.6.1, but acquiring MS/MS on the top 6 multiply charged ions. Phosphopeptides were analysed using multistage activation multistage activation (MSA) (Schroeder et al., 2004), where the Orbitrap performed survey scans at 60,000 resolution, while the LTQ acquired MS/MS on the top 6 multiply charged ions and activated neutral loss species at 98.0, 65.3 and 49.0 *m/z* below the precursor ion mass, each for 30 msec at 35% energy.

Acquired RAW files were searched against the human IPI database (version 3.77), which contained 89,709 entries, using MaxQuant (version 1.2.2.5) and the Andromeda search engine (Cox and Mann, 2008; Cox et al., 2011). For the database search, the minimum peptide length was set to 6 residues and two missed cleavages were permitted. Cysteine carbamidomethylation was set as a fixed modification, while methionine oxidation and S/T/Y phosphorylation were set to variable. Deviations of 7 ppm and 0.5 Th were permitted for precursor and fragment ions, respectively. The false discovery rate was set to 0.1. MaxQuant then further statistically processed the data and performed quantification on SILAC triplet XICs. For quantification, only proteins with at least 3 peptides (one unique) were selected. Each biological replicate was analysed individually and all together in a single iteration of the pipeline.

2.7.5 DATA ANALYSIS

Data obtained from MaxQuant analyses were evaluated with Excel and MeV (version 4.8.1; www.tm4.org/mev). To compare the inter-experimental correlation between proteome and phosphoproteome analysis of each biological replicate, peptides/ phosphopeptides quantified in two or more replicates had their values $\log_{10}$ transformed, plotted on scatter plots and the R^2 correlation(s) visualized as heatmaps. MeV was utilized to perform hierarchical clustering the R^2 data.

Principle component analysis on covariances was performed with JMP10. Peptide data were included for analysis if ratios were available for every Par/ mutant condition. No imputation was performed. Peptides with missing ratios were excluded from the analysis.

GProX was utilised to cluster values output by MaxQuant (Rigbolt et al., 2011b). SILAC ratios were subject to the fuzzy c-means algorithm for unsupervised clustering (Futschik and Carlisle, 2005). 6 clusters were selected, with the fuzzification value set to 2, regulation threshold of ± 0.58 and 200 iterations of the algorithm executed. Clustering was performed together with Dr. Dean Hammond (University of Liverpool).

Gene Ontology (GO) analysis was performed using the DAVID functional annotation tool (version 6.7) (Huang da et al., 2009a), using the GOTERM BP FAT. IPI identifiers of shortlisted proteins and phosphopeptides were input as gene lists, with the background list comprised of all genes identified in the MS analysis as a whole. Threshold count was 2 and EASE score was set to 1. GO terms with a p-value ≤ 0.1 were $\log_{10}$ transformed and visualised with JMP11 as heatmaps.

Only class 1 phosphosites were considered in the phosphoproteome analysis. (localisation probability ≥ 0.75 and a score differential ≥ 5). For linear kinase motif analysis, sequence windows ± 6 residues adjacent to the phosphosite were extracted from MaxQuant output files, before being subject to MotifX analysis (Chou and Schwartz, 2011). Over-represented motifs were identified through comparison with the IPI human proteome (p-value 1×10^{-6}, occurrence limit ≥ 1). Class 1 phosphosites were also submitted to NetworKIN (version 2.0) to identify putative kinase regulators (Linding et al., 2008). Proteins were grouped by MotifX-identified motifs and percentage probability of being regulated by each indicated kinase, before being visualised as a heatmap using MeV. To demonstrate the overall regulation of sites associated with each motif (output), mutant/ Parental ratios from GProX clusters 5 and 6, as well as all phosphosites identified (long list), were normalised by assigned maximum increase and decrease in phosphorylation a value of 1 or -1, respectively. These predictions were performed with the aid of Prof. Ian Prior (University of Liverpool).

3 – RESULTS

3.1 DEVELOPMENT OF AN SRM-BASED RAS QUANTIFICATION TECHNIQUE COUPLED WITH PSAQ

Due to its ability to detect and quantify thousands of peptides/ proteins from biological samples, shotgun proteomics has become a highly popular technique to study complex proteomes over the past decade. However, the shotgun approach is not optimal for the systematic quantification of proteins from complex matrices, as it is stochastic in nature and has limited sensitivity, plus only relative quantification of a sample's most abundant components is commonly achieved (Gallien et al., 2011). Consequently, targeted proteomic approaches have emerged that precisely and quantitatively analyse chosen peptides from complex biological backgrounds. Such techniques are based on selected reaction monitoring (SRM), which was introduced in the late 1970s (Baty and Robinson, 1977; Zakett et al., 1978), and provide highly sensitive and selective quantitative proteomic assays (Anderson and Hunter, 2006). SRM-based approaches typically employ a triple quadrupole mass spectrometer (QqQ), where a precursor ion is selected in the first quadrupole, fragmented in the second and a particular fragment ion selected for detection by the third. The chosen precursor-fragment ion pair, known as a transition, can be altered every few milliseconds to enable the concurrent detection of multiple fragment ions from unlabelled and labelled precursor ions. Unlike the shotgun approach, targeted SRM-based strategies require a minimum amount of information before they can begin: (1) target protein(s) must be known, (2) proteotypic peptides identified, (3) suitable transitions for each peptide selected, (4) instrument parameters optimised, (5) detection and quantification limits determined, (6) peptide abundance and signal response defined and (7) each transition validated (Gallien et al., 2011; Lange et al., 2008). These parameters are defined during assay development, which can be a slow and iterative process. SRM-based techniques are compatible with a variety of stable isotope dilution strategies, e.g. AQUA, QconCAT and protein standard absolute quantification (PSAQ) (Barr et al., 1996; Beynon et al., 2005; Brun et al., 2009; Dupuis et al., 2008; Gygi et al., 1999a). Out of these options, PSAQ enables the most accurate protein quantification, since the stable-isotope labelled, full-length protein standard is added at the beginning of the workflow, meaning it will be subject to the same systematic errors and losses as its endogenous counterpart.

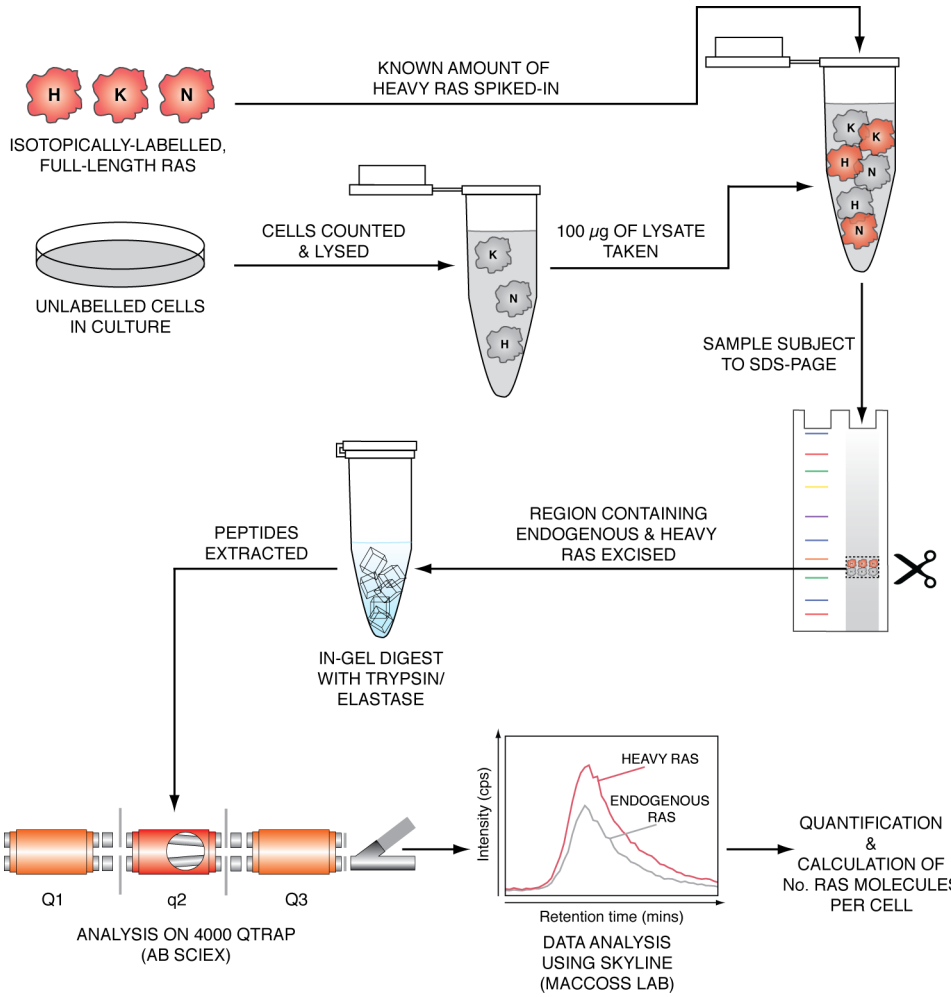


Fig. 3.1. Workflow of the developed SRM-based quantification technique, coupled with protein standard absolute quantification, for measurement of cellular Ras abundance.

Furthermore, because the amino acid sequence of the labelled standard and endogenous protein are identical, digestion of both proteins will be equal. This isn't the case for AQUA and QconCAT strategies since they assume complete digestion and recovery of the endogenous material (Gallien et al., 2011; Lowenthal et al., 2014). This chapter describes the development of an SRM-based proteomic technique, coupled with the PSAQ strategy, for the accurate quantification of cellular Ras abundance (Fig. 3.1). Three human Ras proto-oncogenes encode for four major Ras isoforms: H-, K(A)-, K(B)- and N-Ras. These isoforms are 21 kDa proteins, 188/9 amino acids in length and share complete sequence homology from residues 1-85 and 90% homology from residues 85-165, with the major differences located in the C-terminal hypervariable region (HVR), where there is less than 15% sequence similarity (Karnoub and Weinberg, 2008; Omerovic et al., 2007). The described SRM-based strategy (and PSAQ standards) can distinguish between H-, K(A)-,

K(B)- and N-Ras isoforms and was designed to be utilised in biological laboratories with access to a mass spectrometry facility, since the workflow comprises of methods that are highly familiar to such laboratories or do not require specialist knowledge, e.g. the fractionation of the target lysate is performed using SDS-PAGE and the subsequent in-gel digest step involves simple reduction and alkylation reactions before digestion with well-known and easily obtainable proteases.

3.1.1 IDENTIFICATION OF PROTEOTYPIC RAS PEPTIDES

To establish an SRM-based assay it is necessary to identify peptides that uniquely identify the protein(s) of interest and reliably generate a good MS response. To identify such peptides, known as proteotypic peptides (Mallick et al., 2007), from Ras GTPases is a significant challenge, as the four major isoforms are small proteins that are 85% identical, meaning that a limited number of peptides may be generated by proteolysis and those generated may be shared between all or some isoforms. As a result, four proteases, trypsin, LysC, elastase and GluC, were evaluated for their ability to generate proteotypic peptides that distinguish between each Ras isoform and also provide total Ras information.

3.1.1A TRYPSIN-GENERATED RAS PEPTIDES

Most large-scale proteomic projects utilise trypsin to digest proteins into analysable peptide populations. Trypsin is an aggressive protease, stable under a wide range of conditions and displays high cleavage specificity at arginine and lysine residues (Olsen et al., 2004). Since tryptic peptides contain a C-terminal basic residue, they are typically multiply charged and therefore highly suitable for fragmentation under low-energy CID conditions (Paizs and Suhai, 2005). *In silico* tryptic digestion of Ras isoforms was performed and predicted peptides ≥ 7 ≤ 30 residues in length were searched for in PeptideAtlas (Desiere et al., 2006). Many of the predicted Ras tryptic peptides were contained in the PeptideAtlas MS/MS spectral repository, including LVVVGAGGVK, a peptide that contains positions 12 and 13, and others that would describe each Ras isoform. However, to discover if these peptides were detectable in our hands, 1 μg of his-tagged H-, K(A)-, K(B)- and N-Ras was subject to SDS-PAGE, reduction with dithiothreitol (DTT) and alkylation with iodoacetamide (IAA), before in-gel digestion with trypsin. ~200 ng of peptides were separated by nanoscale C18 RP-HPLC on a nanoAcquity UPLC (Waters) and analysed on an

LTQ Orbitrap XL (Thermo Scientific) using unbiased, data-dependent acquisition of the top 5 most abundant peptides (Fig. 3.2). For all RAW files acquired, each was subject to a Mascot MS/MS ion search (Matrix Science). The detected tryptic Ras peptides, not considering those with missed cleavages, were on average 12.3 ± 4.2 residues long and would describe total wild-type Ras at positions 12 and 13, total K-Ras, and uniquely identify H-Ras, K(A)-Ras, K(B)-Ras and N-Ras. Plus, some detected peptides were shared between two isoforms. To ensure specificity to Ras, all detected peptides, and those described below, were subject to Protein BLAST analysis (Basic Local Alignment Search Tool, National Library of Medicine), using the blastp algorithm and non-redundant protein sequences database (Gish and States, 1993).

3.1.1B LYS-C-GENERATED RAS PEPTIDES

LysC is a protease that cleaves specifically at the C-terminal side of lysine residues (Wada and Kadoya, 2003). *In silico* LysC digest generated a peptide set that was poorly represented in the PeptideAtlas repository. Ras proteins were in-gel digested with LysC and ~200 ng of the subsequent peptides were analysed as described above (Fig 3.3). On average, a LysC peptide was 13.8 ± 5.0 residues in length and those detected could differentiate between H-, K- and N-Ras, but not K(A)- and K(B)-Ras. Furthermore, several of the detected peptides contain mid-chain arginine residues.

3.1.1C ELASTASE-GENERATED RAS PEPTIDES

Elastase is a much less specific protease than trypsin or LysC, with a preference to cleave at the C-terminal side of small neutral amino acids (Rietschel et al., 2009). As a result, it is difficult to predict which Ras peptides will be generated using this enzyme. Ras proteins were digested with elastase before fractionation and analysis as described for trypsin (Fig. 3.4). Many overlapping peptides were observed, yet those detected would enable the differentiation of each Ras isoform and also provide total Ras information. One elastase peptide of interest is FYTLVREI, a pan-Ras peptide, located towards the C-terminal end of Ras, that does not contain residues that are commonly mutated in human cancer (Prior et al., 2012).

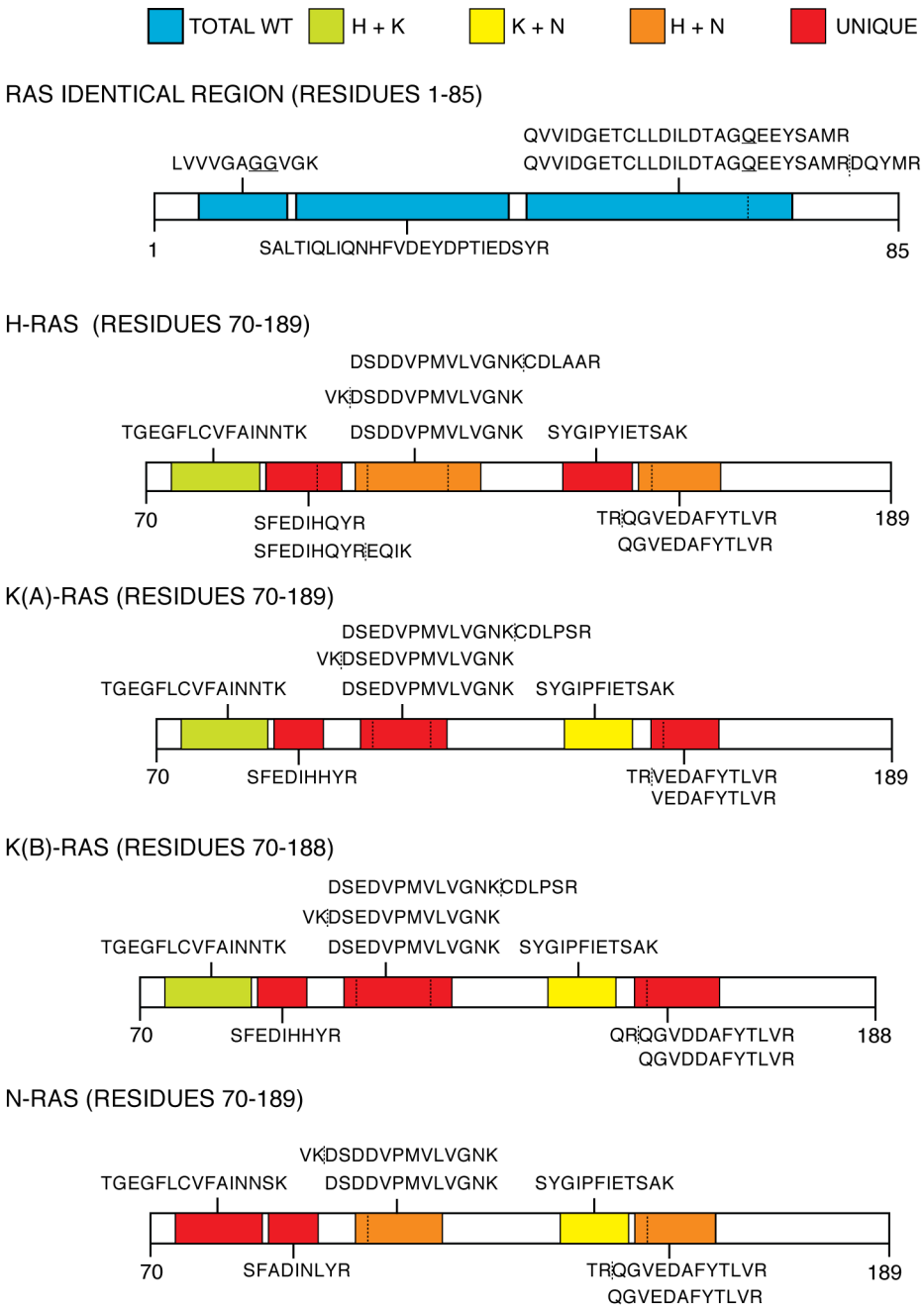


Fig. 3.2. Trypsin-generated Ras peptides. 1 μg of each his-tagged Ras isoform was subject to in-gel digestion with trypsin, before ~200 ng of peptides was analysed by LC-MS/MS on an LTQ Orbitrap XL using data-dependent acquisition of the top 5 most abundant peptides. All peptides were detected in 3 independent runs. Dashed lines indicate sites of missed cleavages. Underlined residues are those commonly mutated in human cancer: G12, G13 and Q61.

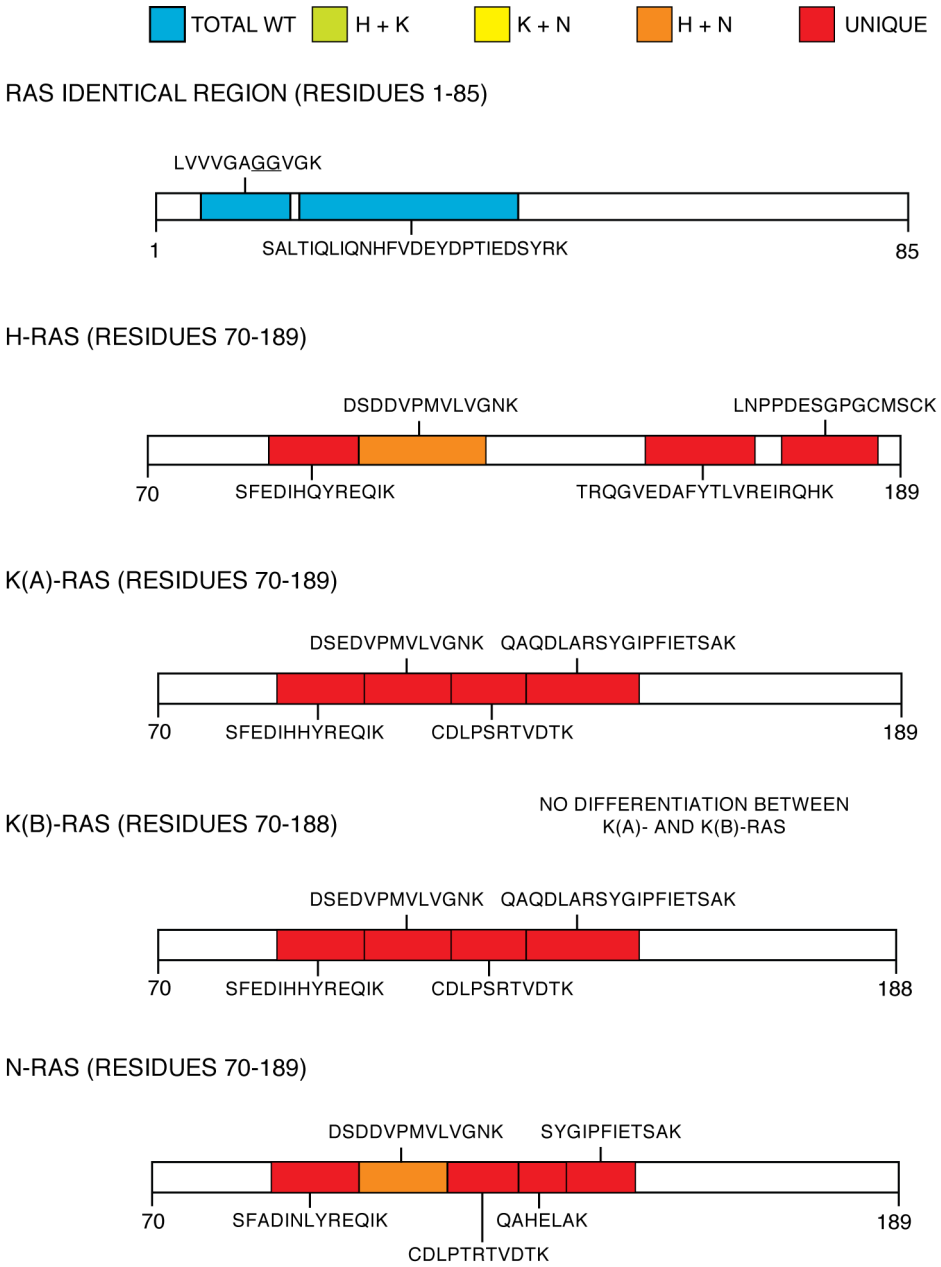


Fig. 3.3. LysC-generated Ras peptides. 1 μg of his-tagged Ras proteins was subject to in-gel digestion with LysC. ~200 ng of peptides were analysed by LC-MS/MS on an LTQ Orbitrap XL using data-dependent acquisition of the top 5 most abundant peptides. All peptides were detected in 3 independent runs. Underlined residues are those commonly mutated in human cancer: G12 and G13.

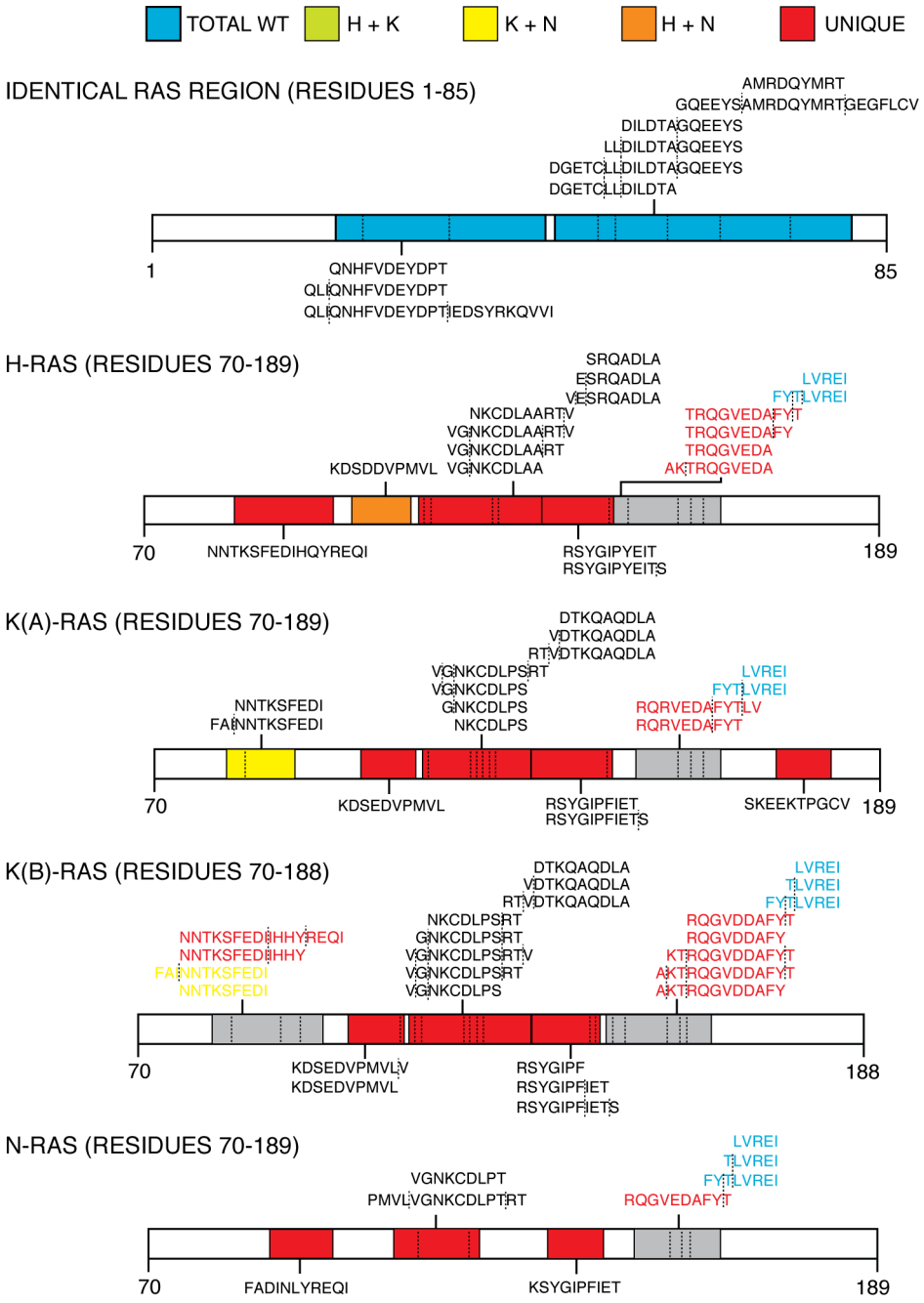


Fig. 3.4. Elastase-generated Ras peptides. ~200 ng of elastase in-gel digested Ras peptides were subject to LC-MS/MS analysis with an LTQ Orbitrap XL using data-dependent acquisition of the top 5 most abundant peptides. Grey regions contain overlapping peptides with different specifies for Ras isoforms. All peptides were detected in 3 independent runs. Dashed lines indicate sites of observed cleavages.

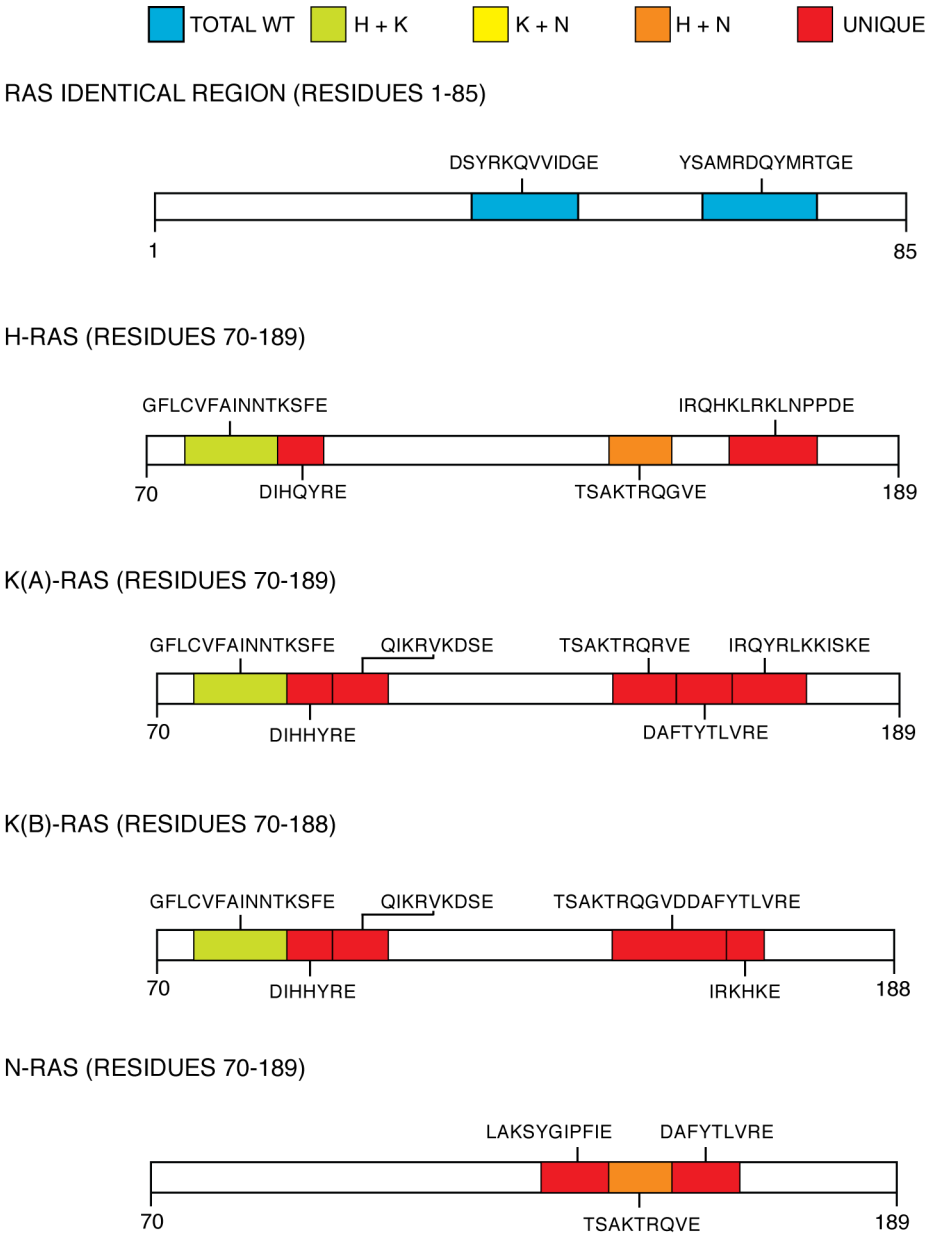


Fig. 3.5. GluC *in silico* digest of Ras isoforms. H-, K(A)-, K(B)- and N-Ras sequences were *in silico* digested with GluC and peptides ≥ 7 ≤ 30 residues in length were collated.

3.1.1D GLU-C-GENERATED RAS PEPTIDES (IN SILICO DIGEST)

The GluC enzyme (*Staphylococcus aureus* Protease V8) specifically cleaves at the C-terminal side of glutamic acid residues, but also cleaves at aspartic acid residues at a 200-300 times slower rate (Birktoft and Breddam, 1994; Drapeau et al., 1972). *In silico* GluC digest was performed (Fig. 3.5) and predicted peptides ≥ 7 ≤ 30 residues in length were input into PeptideAtlas where several were matched to spectra, including DSYRKQVVIDGE, a pan-Ras peptide, and DAFTYLVRE, TSAKTRQGVDDAFYTLVRE and DAFYTLVRE, which are unique for K(A)-, K(B)- and N-Ras, respectively. However, no H-Ras GluC peptide was present in the PeptideAtlas database and many of those present have two or more highly basic mid-chain residues.

3.1.2 PROTEOTYPIC PEPTIDE SELECTION

In a typical SRM-based approach, it is common that only a few representative peptides are used to determine the presence and quantity of a protein in a biological sample (Lange et al., 2008). Consequently, the careful selection of peptides is an essential step in the development of any SRM-based assay. The following criteria was used to determine which peptides were most suitable to measure cellular Ras abundance: (1) the peptide must be unique to Ras, (2) preference will be given to those observed previously, i.e. the peptide is present in a repository, (3) the peptide should not have many chemically modifiable residues, e.g. methionine, glutamine or asparagine, (4) the peptide must not be longer than 22 amino acids, (5) the peptide must not be from the HVR of Ras, (6) where possible, other sites of post-translational modification ought to be avoided, (7) peptides generated due to missed cleavages must be excluded and (8) peptides with multiple mid-chain basic residues should be discounted. An optimal setup is where a single digestion with one protease provides the abundances of each Ras isoform and also that of total Ras. However, none of the tested proteases delivers on this scenario. Yet, from those tested, trypsin generated the largest number of peptides that satisfy the aforementioned criteria and provided the most information regarding Ras abundance from a single digestion (Fig. 3.6).

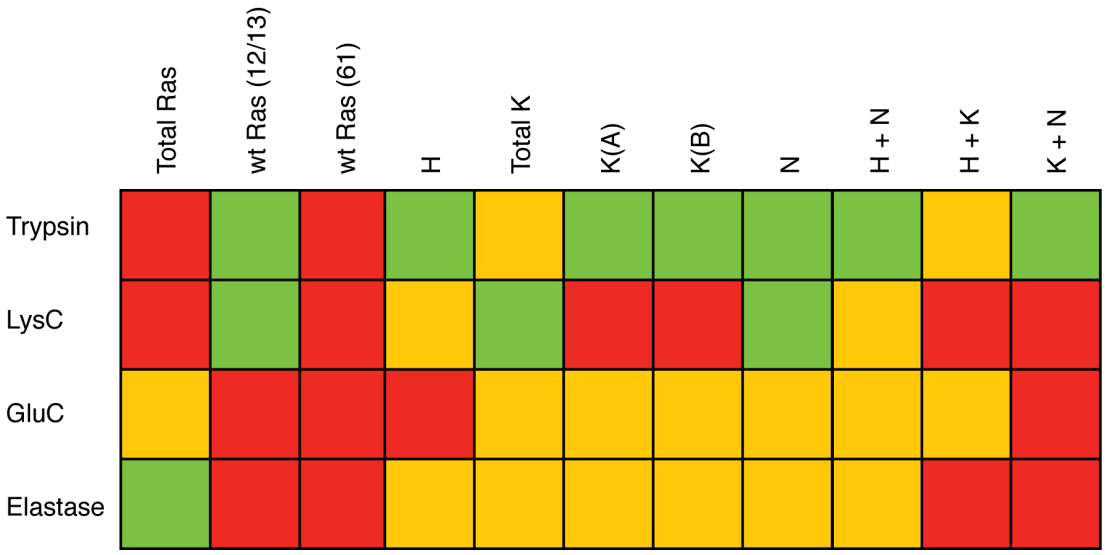


Fig. 3.6. Utility of proteotypic Ras peptides generated using different proteases to describe various aspects of cellular Ras abundance. Green, at least one peptide detected that satisfies criteria stipulated in section 1.2; amber, detected or predicted peptide(s) do not match all criteria; red, peptide not available or those detected/ predicted fail to match most criteria. Total Ras, all Ras isoforms; wt Ras (12/13), wild-type Ras at positions 12 and 13; wt Ras (61), wild-type Ras at position 61; Total K, both K-Ras isoforms. All proteases were tested experimentally, except for GluC.

Trypsin was therefore chosen as the primary protease to obtain information on Ras isoform abundance and total wild-type Ras at positions 12 and 13. Since the two tryptic peptides that describe total Ras are 25 and 26 amino acids long and contain several chemically modifiable residues, they are unsuitable for SRM analysis and were discounted. To complement the trypsin digest, a separate digest with elastase can be performed and the FYTLVREI peptide monitored to obtain total Ras information. Since the LTQ Orbitrap XL is not the instrument platform that was utilised for the SRM assay, it was necessary to detect the candidate proteotypic peptides using a triple quadrupole mass spectrometer. 1 μg of his-tagged Ras isoforms were in-gel digested with trypsin and, in a separate digest, K(B)-Ras was also digested with elastase, subsequently ~200 ng of peptides were subject to LC-MS/MS analysis using information dependent acquisition (IDA) of the top 2 most abundant peptides on a 4000 QTRAP (AB SCIEX). Acquired .wiff files were subject to a Mascot MS/MS ion search. All candidate peptides were detected using this method, except for the elastase peptide FYTLVREI, the K(A)- and K(B)-Ras tryptic peptides VEDAFYTLVR and QGVDDAFYTLVR. MRM-initiated detection and sequencing (MIDAS) (Unwin et al., 2005; Unwin et al., 2009) was employed to detect these peptides instead, since it is more sensitive than IDA (Fig. 3.7).

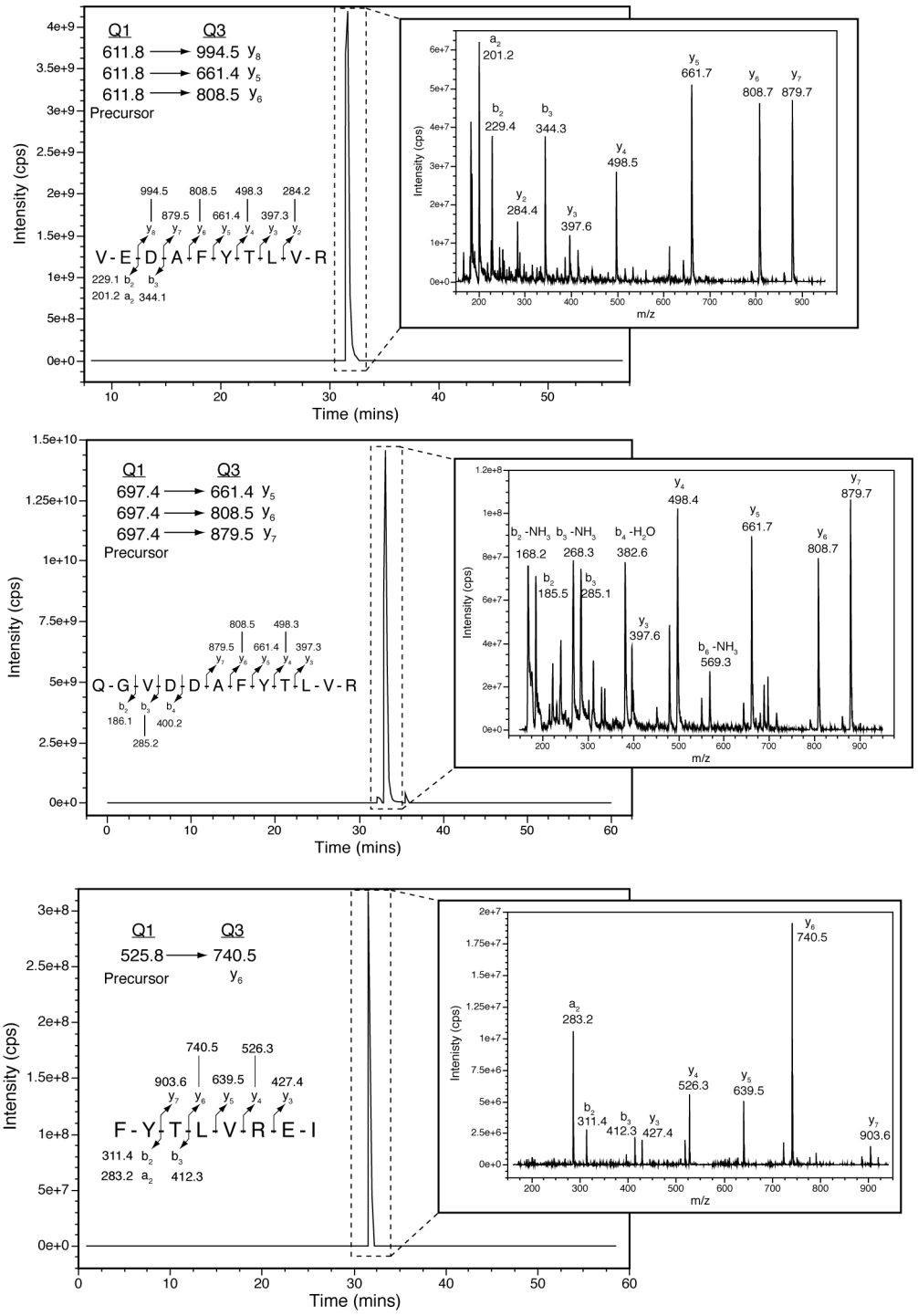


Fig. 3.7. MIDAS detection of proteotypic Ras peptides. MRM-initiated detection and sequencing (MIDAS) (Unwin et al., 2005; Unwin et al., 2009) was performed to detect the indicated peptides on the 4000 QTRAP (AB SCIEX) triple quadrupole mass spectrometer, after they failed to be identified during an initial information dependent acquisition (IDA) scan. Precursor and fragment ion settings are indicated. Graphs show the total ion current and insets display the captured MS/MS spectra.

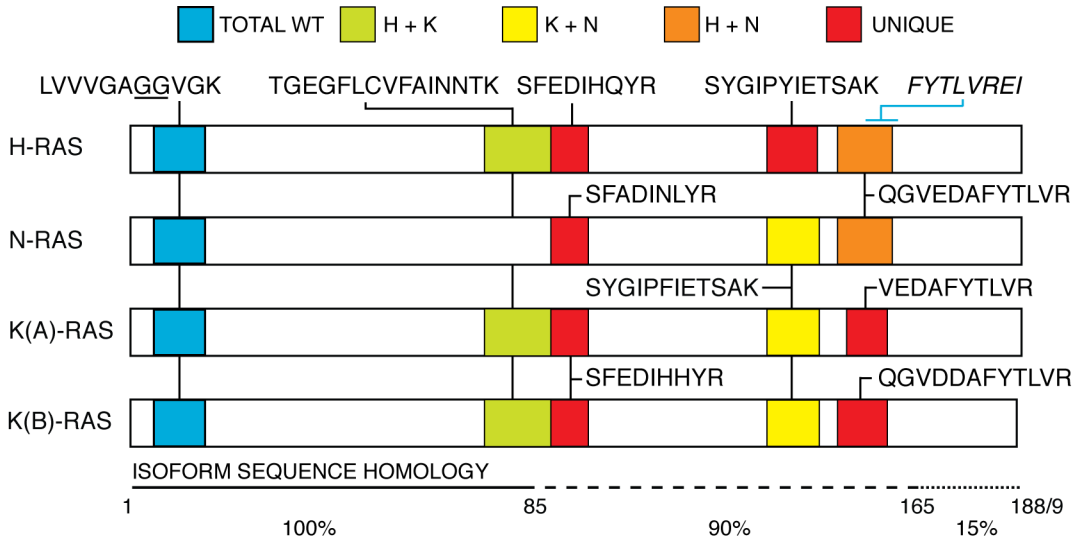


Fig. 3.8. Selected proteotypic Ras peptides. All peptides are tryptic, bar FYTLVREI that is elastase-derived. The candidate peptides describe total Ras, wild-type Ras at positions 12/13, total K-Ras, H-Ras, N-Ras, differentiate between the two K-Ras splice variants and provide redundancy through peptides that are shared between two isoforms. Underlined residues are positions 12 and 13, which are commonly mutated in human cancer.

However, a clear MS/MS spectrum for the total K-Ras peptide DSEDVPMVLVGNK was unattainable when using MIDAS or IDA, meaning this peptide had to be discounted. The selected proteotypic peptide set is shown in Fig. 3.8.

3.1.3 SELECTION OF TRANSITIONS

Peptide quantification using an SRM-based assay requires the selection of specific *m/z* settings for the first and third quadrupole. This combination of *m/z* values will select for a particular precursor/ product ion pair, known as a transition. It is optimal to select between 3 to 5 transitions per peptide, such that a large proportion of the peptide's ion current is measured during the analysis, e.g. 5 transitions will typically account for >70% of the peptide's total ion current (Lange et al., 2008). To obtain the highest sensitivity, it is important to select transitions that correspond to the most intense fragment ions of each precursor, as the intensities of individual fragment ions derived from the same precursor can differ substantially. Based on the acquired MS/MS spectra and following optimisation (see section 1.4), transitions were selected for all candidate Ras peptides that correspond to endogenous/ unlabelled (K0 R0) and isotope-labelled (K8 R10) versions (Table 3.1), as the generated isotope-labelled, full-length Ras protein standards are labelled in this manner (described in section 1.6).

	PEPTIDE	CHARGE	PRECURSOR ION (m/z)		FRAGMENT ION (m/z)		ION	COLLISION ENERGY (eV)	DWELL TIME (msec)
			K0 R0	K8 R10	K0 R0	K8 R10			
Pan	LVVVGAGGVGK	+2	478.3	482.3	743.4	751.4	y9	20	100
	(Wild-type Ras)				644.4	652.4	y8	22	100
					545.3	553.3	y7	24	100
					312.2	312.2	b3	22	100
					185.2	185.2	a2	28	100
	FYTLVREI	+2	520.8	525.8	893.6	903.6	y7	28	200
	(Elastase, total Ras)				730.5	740.5	y6	24	200
					629.5	639.5	y5	24	200
					516.3	526.3	y4	22	200
					283.2	283.2	a2	26	200
H	SFEDIHQYR	+2	597.8	602.8	831.6	841.6	y6	32	200
					716.4	726.4	y5	36	200
					603.3	613.3	y4	32	200
					466.2	476.2	y3	34	200
					207.0	207.0	a2	40	200
	SYGIPYIETSAK	+2	664.8	668.8	908.5	916.5	y8	28	200
					308.1	308.1	b3	34	200
					251.1	251.1	b2	42	200
K	SFEDIHHYR	+2	602.3	607.3	725.5	735.5	y5	38	200
					612.3	622.3	y4	36	200
					475.3	485.3	y3	36	200
					338.2	348.2	y2	38	200
					207.1	207.1	a2	42	200
K(A)	VEDAFYTLVR	+2	606.8	611.8	869.5	879.5	y7	28	200
					798.5	808.5	y6	28	200
					651.4	661.4	y5	26	200
					488.3	498.3	y4	28	200
					229.1	229.1	b2	28	200
K(B)	QGVDDAFYTLVR	+2	692.4	697.4	869.5	879.5	y7	36	200
					798.5	808.5	y6	36	200
					651.4	661.4	y5	36	200
					285.2	285.2	b3	36	200
N	SFADINLYR	+2	549.8	554.8	864.5	874.5	y7	25	200
					678.4	688.4	y5	25	200
					793.4	803.4	y6	27	200
					565.3	575.3	y4	25	200
					207.1	207.1	a2	31	200
H + K	TGEGFL**C**VFAINNTK	+2	835.9	839.9	1066.5	1074.5	y9	38	200
					906.5	914.5	y8	38	200
					807.4	815.4	y7	40	200
					660.4	668.4	y6	38	200
					589.3	597.3	y5	38	200
H + N	QGVEDAFYTLVR	+2	694.5	704.5	869.6	879.6	y7	30	200
					798.6	808.6	y6	30	200
					651.1	661.1	y5	30	200
					488.5	498.5	y4	30	200
					285.2	285.2	b3	30	200
K + N	SYGIPFIETSAK	+2	656.9	660.9	1062.5	1070.5	y10	34	200
					892.4	900.4	y8	28	200
					648.4	656.4	y6	34	200
					421.2	421.2	b4	34	200
					308.1	308.1	b3	34	200

Table. 3.1. Transitions and optimised 4000 QTRAP (AB SCIEX) instrument settings for proteotypic Ras peptides. Transitions were chosen for endogenous/ unlabelled (K0 R0) and isotope-labelled (K8 R10), Ras peptides. K0, isotopically light lysine; R0, isotopically light arginine; K8, $^{13}C_6$ $^{15}N_2$ L-lysine (+8 Da); R10, $^{13}C_6$ $^{15}N_4$ L-arginine (+10 Da). eV, electron volt. Underlined residues indicate the presence of carbamidomethylation.

3.1.4 OPTIMISATION OF INSTRUMENT PARAMETERS

The sensitivity of SRM-based assays is dependent upon proper tuning of instrument parameters. One critical parameter parameter is collision energy (CE), which determines the speed at which ions travel through the q2 collision cell. Default CE values for each peptide were calculated using a generalised equation for doubly charged peptides. CEs ± 2 and ± 4 electron volts (eV) from the default values were tested on all transitions by altering the Q3 *m/z* by 0.01 Da for each CE examined (Sherwood et al., 2009). ~100 ng of Ras peptides (in-gel digested) were separated on a nanoAcquity UPLC (Waters) as described earlier, then subject to analysis on the 4000 QTRAP in SRM mode with 100 msec dwell times for each CE (Fig. 3.9). Over 70% of transitions experienced increases in areas under the curve and peak heights following optimisation, e.g. y9 and y8 ions of LVVVGAGGVGK (Fig. 3.9A), and y8 ion of SYGIPYIETSAK (Fig. 3.9B). Others, like the a2 and y7 ions of LVVVGAGGVGK, however, gained little benefit from the optimisation. As to be expected, the larger Ras peptides required higher collision energies to achieve maximum generation of fragment ions (Fig. 3.9C). All optimised CEs are detailed in Table 3.1.

To enable a robust scheduled SRM assay, a chromatographic gradient was optimised that sufficiently separated the Ras peptides whilst not inducing excessive peak broadening. A 60-minute, 8-35% acetonitrile gradient was selected to separate the Ras peptides using RP-HPLC with a 25 cm × 75 μm (1.7 μm particle size) column and nanoAcquity UPLC system (Waters). This gradient did not cause excessive peak broadening, as most peptides elute over 30 seconds, and nearly all retention times were unique (Table 3.2), with the exception of N- and H-Ras specific peptides, SFADINLYR and SYGIPYIETSAK, respectively, which almost co-elute.

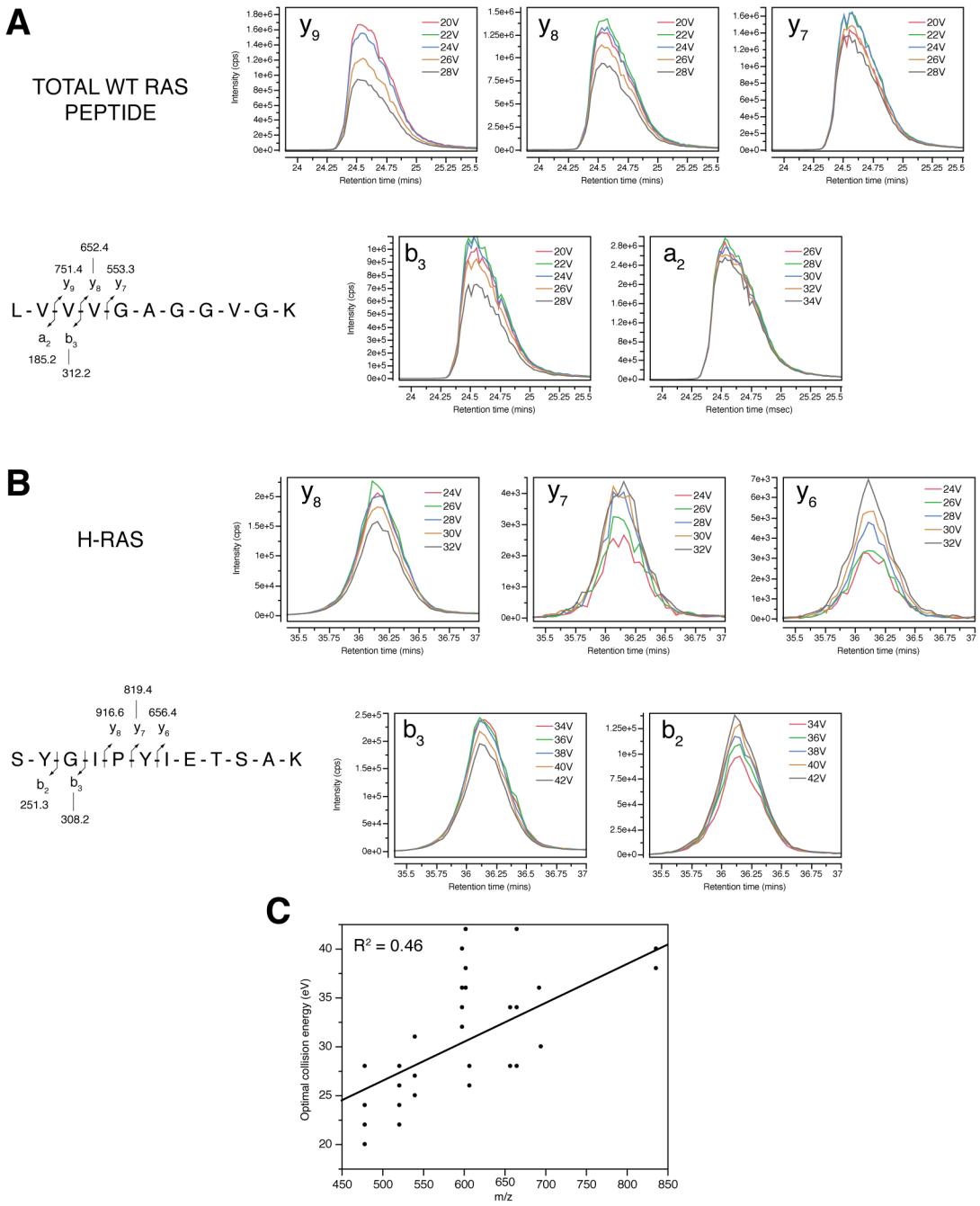


Fig. 3.9. Collision energy optimisation. (**A**) Example collision energy optimisations for LVVVGAGGVGK, a total wild-type Ras peptide, and (**B**) SYGIPYEITSAK, a H-Ras-specific peptide. ~100 ng of Ras peptides were subject to RP-HPLC before analysis on a 4000 QTRAP (AB SCIEX) in SRM mode. Each transition (see Table 3.1) was exposed to ± 2 and ± 4 electron volts from the default value, calculated using a generalised equation for doubly charged ions. (**C**) Overview of optimal collision energies for the proteotypic Ras peptides. Presented data representative of 3 replicates.

	PEPTIDE	PEAK INTENSITY (mins)	PEAK WIDTH (secs)
Pan	LVVVGAGGVGK	30.97 ± 0.37	32.80 ± 0.35
	FYTLVREI (Elastase)	48.38 ± 0.38	24.00 ± 4.20
H	SFEDIHQYR	26.42 ± 0.05	26.20 ± 1.93
	SYGIPYIETSAK	43.69 ± 0.26	41.20 ± 3.62
K	SFEDIHHYR	22.43 ± 0.16	21.60 ± 1.04
K(A)	VEDAFYTLVR	48.21 ± 0.07	25.00 ± 3.30
K(B)	QGVDDAFYTLVR	51.74 ± 0.13	31.00 ± 1.51
N	SFADINLYR	43.23 ± 0.01	26.60 ± 0.35
H + K	TGEGFLCVFAINNTK	61.16 ± 0.01	59.60 ± 1.83
H + N	QGVEDAFYTLVR	55.52 ± 0.03	30.00 ± 0.00
K + N	SYGIPFIETSAK	51.54 ± 0.11	29.20 ± 0.92

Table 3.2. Determination of retention times for proteotypic Ras peptides. ~100 ng of Ras peptides (in-gel digested) were subject to reversed-phase HPLC using a 60-minute, 8-35% acetonitrile gradient on a 25 cm × 75 μm column and nanoAcquity UPLC system (Waters). Peak intensity indicates elution time at which signal intensity was greatest; peak width, time from which the peak reached and returned to 5% maximum intensity. All peptides are tryptic, unless indicated. Data presented as mean ± SD (n = 3).

Dwell times for each transition were chosen to ensure maximum sensitivity whilst obtaining a minimum of 8 data points per transition during the elution of each peptide, a requirement for accurate quantification (Lange et al., 2008). The LVVVGAGGVGK peptide generates the strongest signal and therefore 100 msec dwell times per transition were chosen. For all other peptides, a 200 msec dwell time for each transition was selected, since this limits cycle time to 2 seconds, when 5 light and heavy transitions are monitored, and will therefore obtain ~15 data points for each transition for a peptide that elutes over 30 seconds. Yet, it is very likely that the observed intensities of the Ras peptides will reduce when analysed in the presence of a complex biological background due to ion suppression. Consequently, the 200 msec dwell times should allow for a drop in signal and still maintain a minimum of 8 data points for each transition.

3.1.5 VALIDATION OF TRANSITIONS

Although SRM analyses are highly selective due to the two consecutive mass filtering stages, a particular transition may not be specific in the presence of a complex biological background. Since no MS/MS spectra are acquired during an SRM assay, interfering signals can be easily mistaken as being derived from the target peptide. Several validation steps for each transition were therefore performed. All chosen transitions were based on observed fragment ions from MS/MS spectra obtained using pure Ras protein digests on the 4000 QTRAP in IDA or MIDAS modes. This ensures that all selected transitions are genuine. Further validation steps are performed on the acquired SRM data itself, which are detailed in the following chapter. Briefly, heavy isotope-labelled peptides (with ^{15}N and ^{13}C incorporated) exactly co-elute with their endogenous counterparts. This enables the clear identification of the target peptides from any non-specific signals. Additionally, the signal intensity ratios of the transitions should be identical between the light and heavy peptides. Any change in these ratios would indicate interference with one or more transitions.

Interference between transitions was also probed. Since unlabelled and isotope-labelled versions of the same peptide have identical b-ion series, there is the possibility that cross talk may occur if the collision cell is insufficiently cleared between transitions when b ions are targeted (Fig. 3.10A). ~100 ng of Ras peptides were subject to SRM analysis where LVVVGAGGVGK and VEDAFYTLVR peptides were monitored and a b-ion was selected for transmission (true transition). Immediately after this transition, another was acquired that selects for the same b ion in Q3 but a non-existent precursor in Q1 (false transition) (Fig. 3.10B). There was no detectable level of cross talk in the false transition, which therefore permits the use of b ions in the SRM assay.

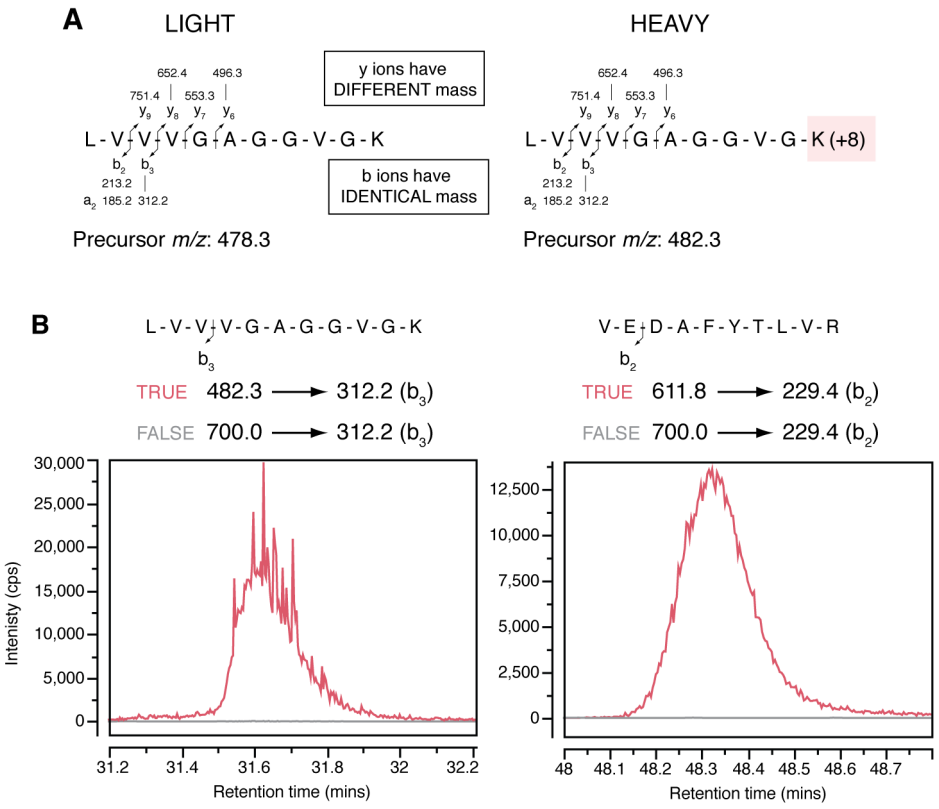


Fig. 3.10. Cross talk between transitions targeting b ions on the 4000 QTRAP. (**A**) Although unlabelled and isotope-labelled peptides have different precursor and y ion masses, their b ion series are identical. (**B**) Determination of no cross talk between transitions. ~100 ng of Ras peptides were subject to SRM analysis on the 4000 QTRAP monitoring for LVVVGAGGVGK and VEDAFYTLVR in Q1 and a b-ion in Q3 (true transition). This was immediately followed by a transition that selects for the same Q3 mass, but a Q1 mass that selects for a non-existent precursor (false transition).

Furthermore, to determine that the SDS-PAGE fractionation technique, selected transitions and optimised instrument settings in combination are compatible with generating a linear relationship between Ras peptide abundance and MS response on the QqQ, 250 ng of unlabelled Ras protein isoforms were mixed with isotope-labelled Ras protein at ratios from 0.2 - 20, before being subject to SDS-PAGE and in-gel digestion. Generation of isotope-labelled, full-length Ras protein standards is detailed in the next section. 50 ng of light Ras peptides, and between 10 ng and 1 μg of isotope-labelled peptides, were injected onto a RP-HPLC column, using the optimised gradient described previously, before SRM analysis using the transitions and settings stipulated in Table 3.1. All tested Ras peptides demonstrated a linear response between peptide abundance and signal intensity, with R^2 values between 0.999 and 0.978 and slopes all close to 1 (Fig. 3.11).

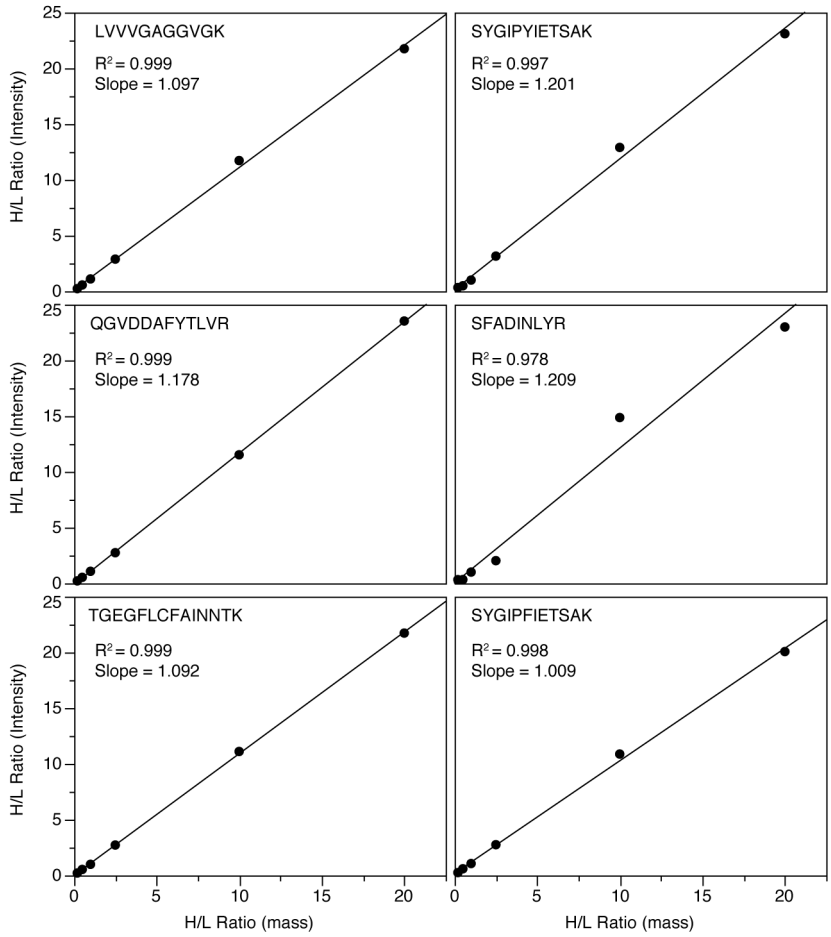


Fig. 3.11. Linearity between Ras peptide abundance and MS response. Isotopically light, full-length Ras isoforms were mixed with isotope-labelled Ras proteins at ratios from 0.2 to 20, before being subject to in-gel digestion. 50 ng of light Ras peptides, and between 10 ng and 1 μg of heavy Ras peptides, were injected onto a RP-HPLC column and subject to SRM analysis on a 4000 QTRAP using the transitions and settings stipulated in Table 3.1.

At 10 and 25 ng, only 2 clear transitions were detected for the K-Ras total peptide, SFEDIHHYR, and H-Ras specific peptide, SFEDIHQR. Both of these peptides were detected in doubly and triply charged states, plus several of their fragment ions were either singularly or doubly charged. Also, if the y4 ion of the K(B)-Ras specific peptide, QGVDDAFYTLVR, is monitored along with the 4 selected transitions, R^2 is maintained at 0.999 but the slope increases to 1.374. As a result, the y4 transition does not behave in the same way as the other selected transitions and therefore cannot be utilised in the SRM assay.

3.1.6 PRODUCTION OF PSAQ RAS STANDARDS

Since the PSAQ strategy requires the generation of full-length, isotope-labelled protein standards and the selected proteotypic Ras peptide set contain either a C-terminal arginine or lysine residue (mid-chain arginine for the selected elastase peptide), full-length, his-tagged Ras protein isoforms were produced labelled with $^{13}C_6$ $^{15}N_2$ L-lysine (+8 Da) and $^{13}C_6$ $^{15}N_4$ L-arginine (+10 Da). To express these proteins, auxotrophic AT713 *E.coli* (Yale Coli Genetic Stock Center, New Haven) were used. These bacteria are deficient for ArgA and LysA biosynthetic genes (*argA21, lysA22*), making them suitable for generating protein uniformly labelled with arginine and lysine isotopes.

Wild-type H- and K(A)-Ras sequences were cloned into the pTrcHis A vector (Invitrogen), while pRSET vectors (Invitrogen) containing his-tagged, wild-type K(B)- and N-Ras sequences were kindly provided by Ignacio Rubio (Institute of Molecular Cell Biology, University of Jena). AT713 *E.coli* were transformed with these plasmids and grown in M9 minimal media, supplemented with heavy arginine and lysine (Fig. 3.12). The his-tagged Ras isoforms were initially purified using a USB PrepEase his-tagged protein purification kit (Affymetrix), before being subject to gel filtration using a Superdex 200 10/300 GL and ÄKTA protein purification system (both GE Healthcare). The Ras PSAQ standards were not contaminated with any other protein and over 99% isotopically labelled (Fig. 3.13). Concentrations of the Ras PSAQ standards were determined by generating an isotopically light K(B)-Ras protein, whose concentration was measured using two independent methods: absorbance measurement at 280 nm using a NanoDrop 1000 and a bicinchoninic assay (both Thermo). The two methods gave results that were only 9% different. A known amount of the light K(B)-Ras protein was mixed with volumes of the Ras PSAQ standard, before in-gel digestion and SRM analysis of light and heavy versions of the shared LVVVGAGGVGK peptide on the 4000 QTRAP. Comparison of the intensities of the respective isotope signals enabled accurate measurement of the PSAQ standards (Fig. 3.14).

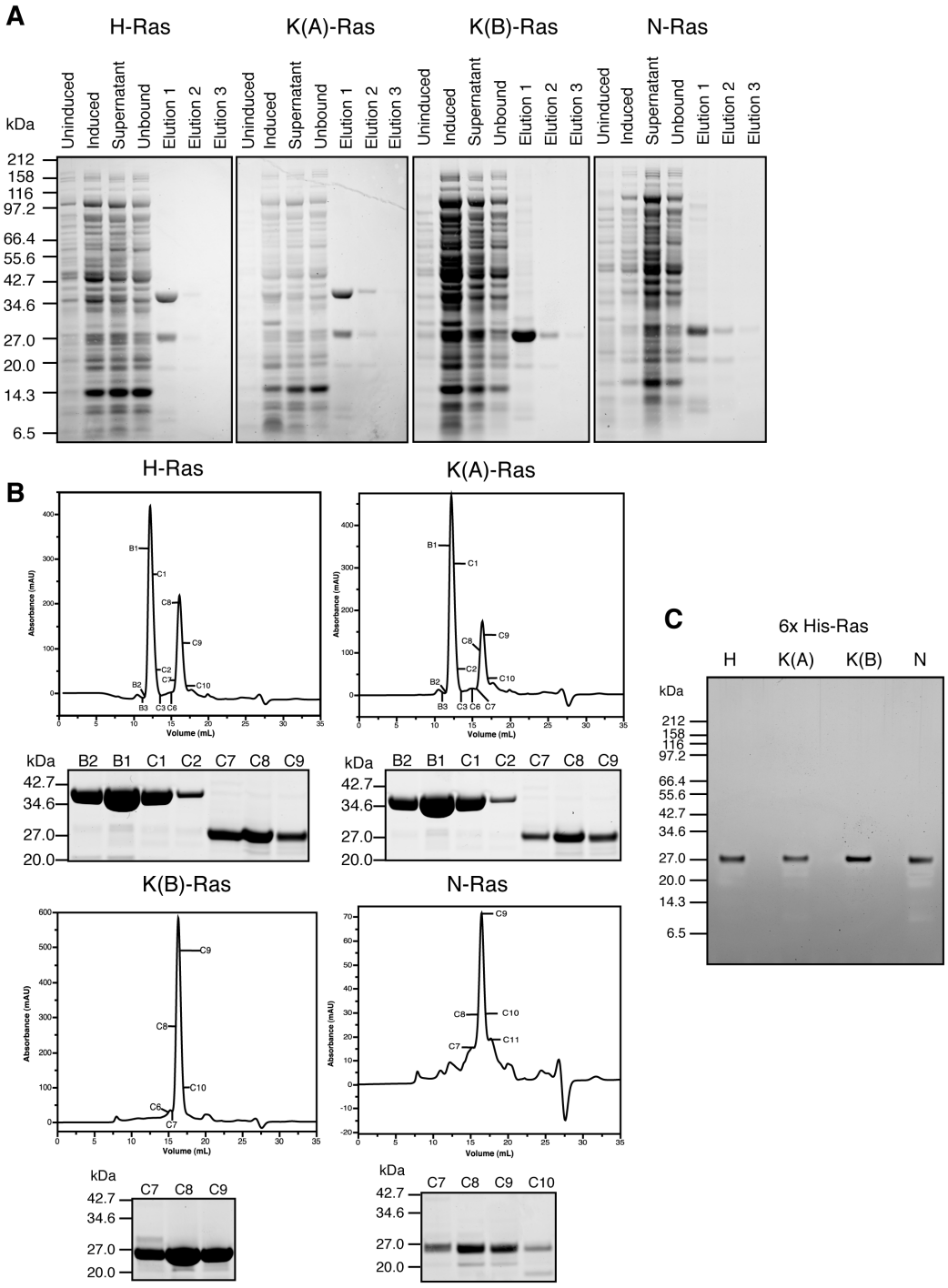


Fig. 3.12. Production of isotope-labelled (K8, R10) H-, K(A)-, K(B)- and N-Ras protein isoforms. (**A**) Expression and initial purification of isotope-labelled Ras proteins. Auxotrophic AT713 *E.coli* were transformed with plasmids containing his-tagged, wild-type Ras isoform sequences and grown in isotopic M9 minimal media, containing heavy arginine and lysine, for 3 hours following induction with isopropyl β-D-1-thiogalactopyranoside. Bacteria were pelleted by centrifugation and his-tagged proteins purified using a PrepEase purification kit (Affymetrix). (**B**) Further purification of his-tagged Ras isoforms using gel filtration. The first elution of each protein was further purified using size exclusion chromatography with a Superdex 200 GL (GE healthcare). UV traces and colloidal blue staining of 25 μL of the indicated fractions are shown. (**C**) Final purity of the his-tagged Ras protein standards. Fractions C7-C9 were combined for all Ras isoforms, bar N-Ras where only the C7 fraction was collected. The band at 34 kDa in H- and K(A)-Ras gels represents a bacterial protein contaminant, which was removed via the gel filtration.

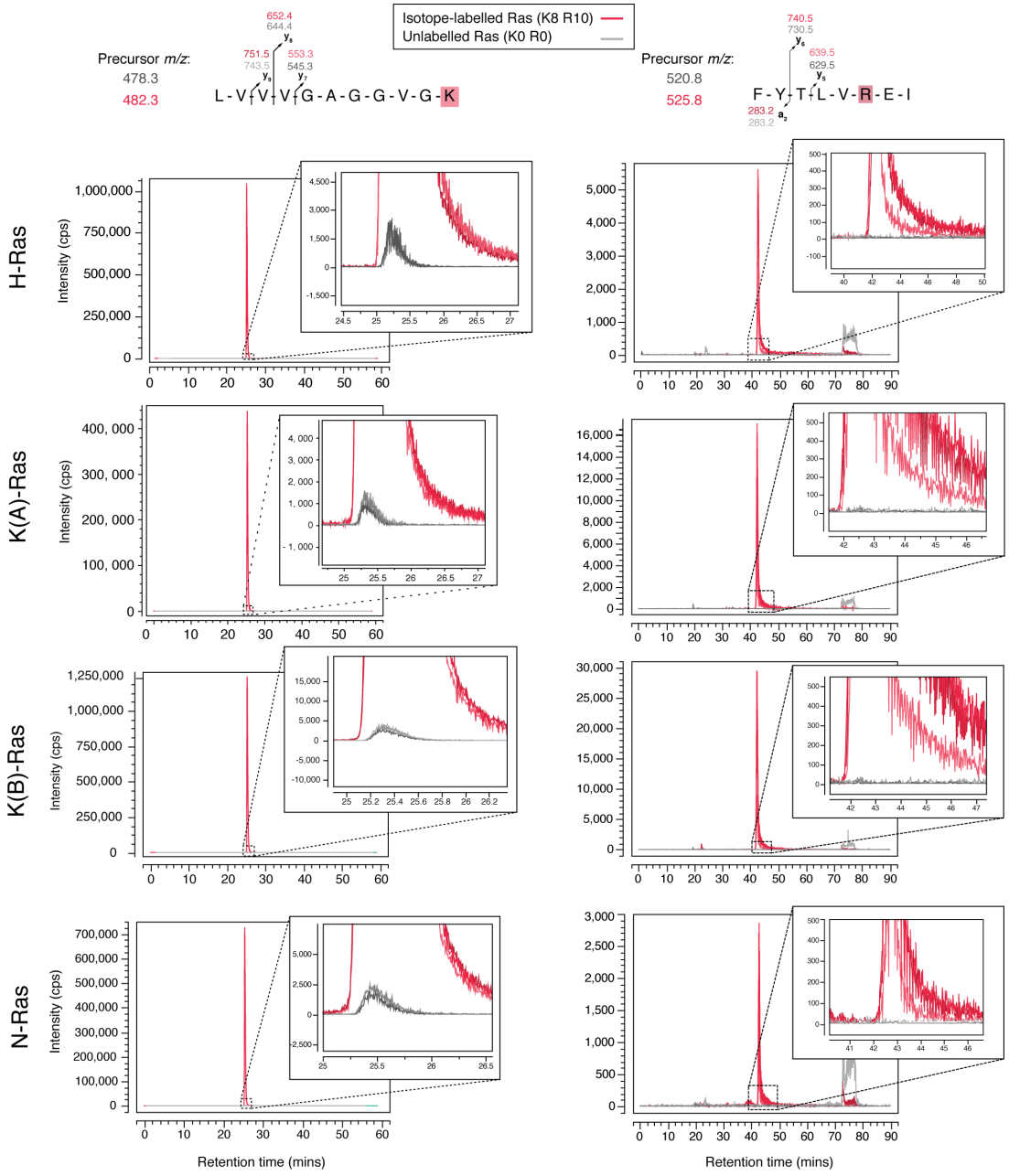


Fig. 3.13. Extent of isotope labelling of the full-length, PSAQ Ras isoform standards. Each his-tagged, isotope-labelled (K8 R10) Ras protein standard was subject to separate in-gel digestions with trypsin and elastase, before analysis of light and heavy versions of LVVVGAGGVGK and FYTLVREI peptides using a 4000 QTRAP (AB SCIEX) in SRM mode. The standards were > 99% labelled on lysine residues, whilst no light version of the arginine containing peptide was detected.

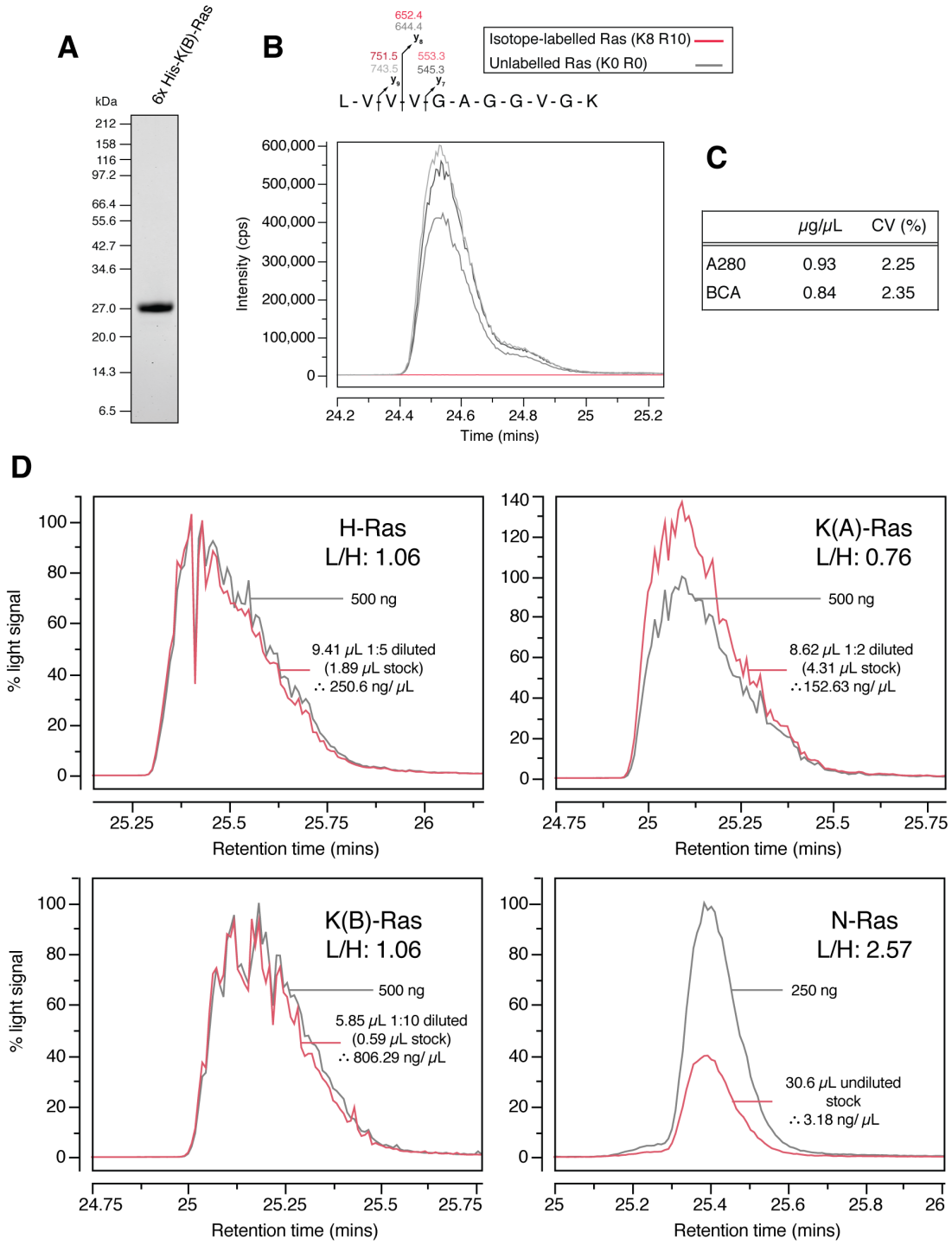


Fig. 3.14. Concentration of the isotope-labelled (K8 R10), Ras PSAQ standards. (**A**) Colloidal blue staining of 1μg and (**B**) total light labelling of full-length, his-tagged K(B)-Ras. (**C**) Concentration of the light His-K(B)-Ras protein. Two independent methods were utilised: absorbance at 280 nm (A280) and bicinchoninic assay (BCA). Performed in triplicate. Mass of the his-tag (17% of total protein) was deducted. (**D**) Concentration of heavy-labelled H-, K(A)-, K(B)- and N-Ras proteins. A known amount of the accurately quantified light His-K(B)-Ras (250 or 500 ng) was mixed with volumes of isotope-labelled (K8 R10) PSAQ Ras standards and subject to in-gel tryptic digestion, before RP-HPLC and SRM analysis on a 4000 QTRAP, monitoring for light and heavy forms of the LVVVGAGGVGK peptide (transitions described in Table 3.1). Comparing the relative intensities of the different isotope signals enabled accurate measurement of the Ras PSAQ standards.

3.1.7 DISCUSSION

Since the four major Ras isoforms are 85% identical and only 188/9 residues in length, they present a significant challenge to be targeted by SRM-based proteomics. As a result, four proteases (trypsin, LysC, GluC and elastase) were evaluated for their ability to generate proteotypic Ras peptides that distinguish the Ras isoforms. Candidate peptides were selected based on their amino acid sequence, likely charge state(s), presence of chemically modifiable residues and known Ras PTMs. From the proteases tested, trypsin generated the largest number of proteotypic peptides that were amenable to SRM analysis and defined the most aspects of cellular Ras abundance. Trypsin differentiates the four major isoforms, describes wild-type Ras at positions 12/ 13 and provides redundancy with peptides shared between isoforms. Yet, a tryptic peptide that defines total Ras levels was lacking, as the Ras identical region (residues 1-85) contains only 6 arginine and lysine residues. Consequently, an elastase-derived peptide FYTLVREI (residues 156-163) may be utilised to obtain total Ras information, but would require a separate digestion.

Tryptic peptides are commonly observed as doubly charged ions, with one proton sequestered by their C-terminal basic residue whilst the other, following excitation, can migrate to energetically less favourable sites along the peptide backbone to cause dissociation (Paizs and Suhai, 2005). As the majority of selected tryptic Ras peptides do not contain a highly basic amino acid other than at their C-terminus, they are likely to reside in a single charge state. Yet, two tryptic Ras peptides contain mid-chain histidine residues: SFEDIHHYR (K-Ras) and SFEDIHQYR (H-Ras). These were detected as doubly and triply charged precursor ions that fragment into a mix of singly and doubly charged product ions. This is not ideal for SRM, as the mid-chain histidine residues may sequester the additional proton, making the peptide less amenable to fragmentation by CID. Moreover, the signal from these peptides is spread across multiple charge states. Both of these factors will reduce the sensitivity of the SRM assay, possibly explaining why these peptides were inadequately detected from in-gel digests of 10 and 25 ng of pure Ras protein. Overall, this potentially jeopardises the detection of these peptides at endogenous levels. However, trypsin provides alternative H- and K-Ras peptides that can be utilised if these peptides are not detectable in a target lysate.

The presence of proline residues in peptides can be advantageous in terms of sensitivity for SRM, as peptides preferentially fragment N-terminally of these residues and generate an intense y-ion signal (Schwartz and Bursey, 1992). Two selected tryptic Ras peptides contain proline residues, the H-Ras specific peptide SYGIPYIETSAK and the shared H- and N-Ras peptide, SYGIPFIETSAK. Both generated a dominant y8 ion signal, exemplified by the H-Ras peptide, where only two other fragment ions, b2 and b3, were viable choices for transitions. y7 and y6 product ions were initially chosen, but following CE optimisation, their intensities remained ~100-fold less than that for the y8 ion. As a result, this peptide is quantified using only 3 transitions.

Peptides that contain methionine, asparagine and glutamine can be chemically modified and therefore present problems for SRM-based strategies. When modified, these residues experience a change in mass. Methionine can undergo oxidation, generating methionine sulfoxide or sulfone, causing a mass increase of 15.99 or 31.99 Da, respectively (Ghesquiere and Gevaert, 2014). Asparagine and glutamine residues can undergo deamidation and be converted into isoaspartate/ aspartate and glutamate, respectively (Geiger and Clarke, 1987; Lange et al., 2008; Piszkiewicz et al., 1970). This modification causes an increase in mass of almost 1 Da, but occurs at a much lower rate for glutamine residues. Such mass changes are important, because Q1 will eject modified peptides if not specifically selected for in SRM. Chemical modifications, like different charge states, will also act to spread a peptide's signal and reduce the sensitivity of the assay. Tryptic peptides with methionine residues were therefore discounted: DSEDVPMVLVGNK (K-Ras) and DSDDVPMVLVGNK (H- and N-Ras). As tryptophan can also be oxidised, peptides containing this residue would also be discounted, but interestingly Ras does not contain this residue. Since Ras is a small protein, it was not possible to avoid peptides that contain asparagine or glutamine: TGEGFLCVFAINNTK (H and K), QGVEDAFYTLVR (H and N) and QGVDDAFYTLVR (K(B)-Ras). Although these peptides are prone to chemical modification, the use of the PSAQ strategy ought to compensate for acquired modifications, since the heavy standard is added at the earliest possible point in the workflow, both the isotope-labelled and endogenous proteins should be equally affected by such modifications.

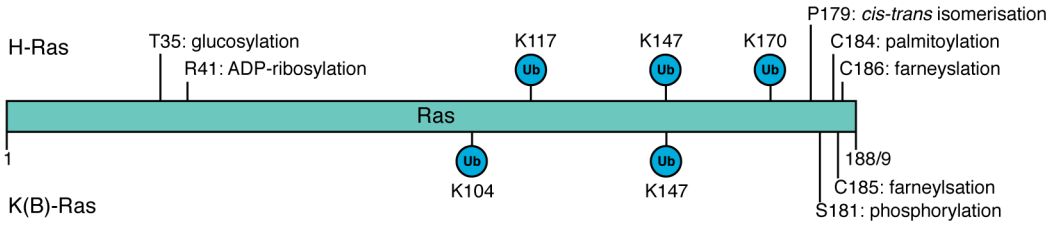


Fig. 3.15. Post-translational modifications of Ras. Known modifications of H-Ras (top) and K(B)-Ras (bottom). Based on (Ahearn et al., 2012).

The only point of variation would be if there are differences between the endogenous protein isolated from cells and the isotope-labelled standards, but this is largely unavoidable. Yet, non-acidic buffers were used during the purification of isotope-labelled standards and cell lysis to limit the extent of deamidation.

Ras PTMs must also be considered when selecting peptides for quantification, as modified peptides must be specifically selected if they are to be detected using SRM. While K(B)-Ras requires farnesylation to associate with the plasma membrane, the remaining isoforms require further palmitoylation to achieve membrane binding (Ahearn et al., 2012; Prior and Hancock, 2012). These lipid modifications occur in the HVR of Ras and therefore any peptide in this region was discounted. However, Ras proteins can also be modified on residues outside of this region (Fig. 3.15). From the selected peptides (Fig. 3.8), ubiquitylation of K147 will affect SYGIPYIETSAK (H-Ras) and SYGIPFIETSAK (K- and N-Ras) (Jura et al., 2006; Sasaki et al., 2011). However, the detected levels of ubiquitylated H-Ras were between 1-2% of total H-Ras, meaning this peptide should be capable of reporting accurate H-Ras levels despite being modified. Yet, the extent of K- and N-Ras ubiquitylation at K147 is unclear, but the use of this peptide can be validated if individual values for cellular K- and N-Ras match its results.

LysC, GluC and elastase were not chosen as the primary protease to acquire information on cellular Ras abundance. LysC and GluC generated peptides with multiple mid-chain highly basic residues, meaning the fragmentation of these peptides may be limited and, as described earlier, their signals will be diluted. Furthermore, LysC generated longer peptides compared to trypsin and most tended to contain chemically modifiable residues. Elastase generated the most peptides compared to the other enzymes tested, but many were overlapping. When peptides

overlap, their signal is shared and the possibility of detecting them in a target lysate is reduced. Consequently, these proteases were only considered to complement the tryptic digest. Elastase digestion consistently generated the pan-Ras peptide FYTLVREI, while GluC generated the pan-Ras peptide DSYRKQVVIDGE. Neither of these peptides is affected by common Ras mutations and therefore suitable to define total cellular Ras. These proteases can therefore be utilised to obtain total Ras information in a separate digest alongside trypsin.

Once the Ras proteotypic peptides were chosen, suitable transitions for each peptide were selected. To ensure that each transition was genuine, only fragment ions observed in MS/MS spectra obtained using the 4000 QTRAP were considered. y-ions were given preference over b- and a-ions, since y-ions contain more of the peptide they will have a more unique mass. Where possible, a minimum of three y-ions was selected per peptide, supplemented with up to two a- or b-ions. Transitions were optimised with respect to CE in order to increase the intensity of each fragment ion and improve assay sensitivity. Following optimisation, unlabelled and isotope-labelled proteins were mixed at various ratios and subject to in-gel digestion and SRM analysis with the optimised instrument settings. All tested peptides demonstrated a linear relationship between Ras peptide abundance and signal intensity. However, the K(B)-Ras peptide, QGVDDAFYTLVR, demonstrated an interesting phenomenon, where the inclusion of a y4 ion caused the slope of the MS response to increase. This indicates that, at higher abundances the isotope-labelled y4 ion is more intense than its unlabelled counterpart. This is puzzling, but ultimately means the y4 ion cannot be employed in the SRM assay. Along with optimised instrument parameters, a 60-minute RP-HPLC linear gradient was designed. All peptides displayed unique retention times, bar, SFADINLYR and SYGIPFIETSAK, which nearly co-elute. This unfortunately necessitates the use of two independent LC runs to monitor both of these peptides when using the 4000 QTRAP, but this shouldn't be necessary, faster and more sensitive on newer QqQ instruments. The developed assay was readily capable of detecting all Ras isoforms at 10 ng levels and is compatible with samples derived from human, mouse, rat, chicken and guinea pig tissues. In the next chapter, the SRM Ras quantification technique was applied to quantitate Ras isoform abundance in isogenic SW48 colorectal cancer cell lines, harbouring a variety of K-Ras codon mutations, as well as H-Ras and N-RasG12V mutations.

3.2 QUANTIFICATION OF CELLULAR RAS ABUNDANCE

Ras GTPases are small, predominately plasma membrane-localised proteins, which sit near the top of signalling pathways regulating cell proliferation, differentiation and apoptosis (Prior and Hancock, 2012). All human cells harbour three Ras proto-oncogenes, that encode for four major protein isoforms: H-, K(A)-, K(B)- and N-Ras (Fiorucci and Hall, 1988; Furth et al., 1987), and are amongst the most frequently mutated in human cancer, with an average mutation incidence rate of 16% (Prior et al., 2012). Although Ras isoforms share the same effector- and nucleotide-binding regions, and are overall 85% identical, they are not functionally redundant (Castellano and Santos, 2011). Despite being studied for over 30 years, a complete understanding of Ras function and dysfunction, particularly in human cancer, is yet to be realised. Moreover, it is still uncertain how many Ras molecules are present in a cell. This is a critical parameter, especially when cell signalling research is nearing the end of identifying/ characterising individual proteins, but shifting from a reductionist approach to a more holistic one to understand how proteins work together to form complex molecular networks and circuits using mathematical and physicochemical modelling (Aldridge et al., 2006; Bordbar et al., 2014).

Several studies have attempted to quantify cellular Ras abundance in recent years; however, their findings are conflicting. Omerovic et al. utilised an immunoblotting strategy to show that K-Ras is relatively the most abundant isoform in several commonly used cell lines (Omerovic et al., 2008). In contrast, Lampson et al. suggest that K-Ras is the minor isoform, due to ribosomal stalling on rare codons within in the KRAS gene to cause poor K-Ras protein expression; plus, they demonstrated that H-Ras was ~20-fold better expressed than K-Ras when identically tagged cDNA was transfected into cells (Lampson et al., 2013). SRM-based proteomics has also been applied to the question of cellular Ras abundance. Wang et al. performed immunoprecipitation of Ras proteins, before tryptic digestion and addition of AQUA reference peptides, ahead of SRM analysis (Wang et al., 2011). The Wang study found that, on average, there are over 1 million Ras molecules per cell, with K-Ras the most abundant isoform. Confusingly, Ruppen-Canas et al. employed the same technique as the Wang study, but detected ~100-fold less Ras in their samples (Ruppen-Canas et al., 2012). Using a different approach, Halvey et al. subjected samples to SDS-PAGE and in-gel digestion,

before addition of Ras-specific AQUA peptides and SRM analysis (Halvey et al., 2012). They found a ~10-fold less Ras in their samples compared the Wang study, but could not convert this into molecules per cell. Global, label-free proteomic studies also provide a number for Ras molecules per cell. Schwanhausser et al. and Nagaraj et al. both employed the intensity based absolute quantification (iBAQ) technique to measure the abundances of thousands of proteins in mouse NIH3T3 and HeLa cells, respectively (Nagaraj et al., 2011; Schwanhausser et al., 2011). The Nagaraj study quantified 173,933 K-Ras molecules per HeLa cell, while the Schwanhausser study calculated 22,089 H-Ras, 41,089 K-Ras and 378,351 N-Ras molecules per NIH3T3 mouse cell. Overall, there have been a wide range of numbers reported for Ras molecules per cell and disagreement over which isoform is most abundant. Here we measure the cellular Ras abundance in a panel of isogenic SW48 colorectal cell lines, harbouring a variety of Ras mutations, utilising the SRM-based Ras quantification strategy and PSAQ standards described in the previous chapter. In addition, we investigate the turnover of wild-type and mutant Ras molecules utilising the SW48 isogenic cell line panel, coupled with dynamic SILAC.

3.2.1 IMMUNOBLOT ANALYSIS OF RAS ISOFORM ABUNDANCE

Firstly, we assessed the utility of immunoblotting to quantify cellular Ras abundance, since this is a well-established technique to identify specific proteins and determine their expression levels in tissues or cells in culture. 100 to 10 ng of his-tagged H-, K(B)- and N-Ras, along with 25, 50 and 100 μg of HeLa cell lysate, was subject to immunoblotting with H-, K-, N- and Pan-Ras antibodies (Fig 3.16A). Densitometry analysis of the his-tagged Ras proteins provides a standard curve from which Ras isoform abundance in the HeLa lysates can be calculated (Fig. 3.16B). Unlike the K- and pan-Ras antibodies, which displayed largely linear responses between his-tagged Ras abundance and densitometry signal in all replicates, the H-Ras antibody failed to exhibit a linear response, while the N-Ras antibody achieved such a response only once. When Ras abundance in HeLa cells was calculated, the variability between biological replicates was high, e.g. in 50 μg lysate, H-Ras was calculated as 15.39 ± 8.32 ng, K-Ras 8.20 ± 9.88 ng and N-Ras 10.88 ± 8.31 ng (Fig. 3.2C).

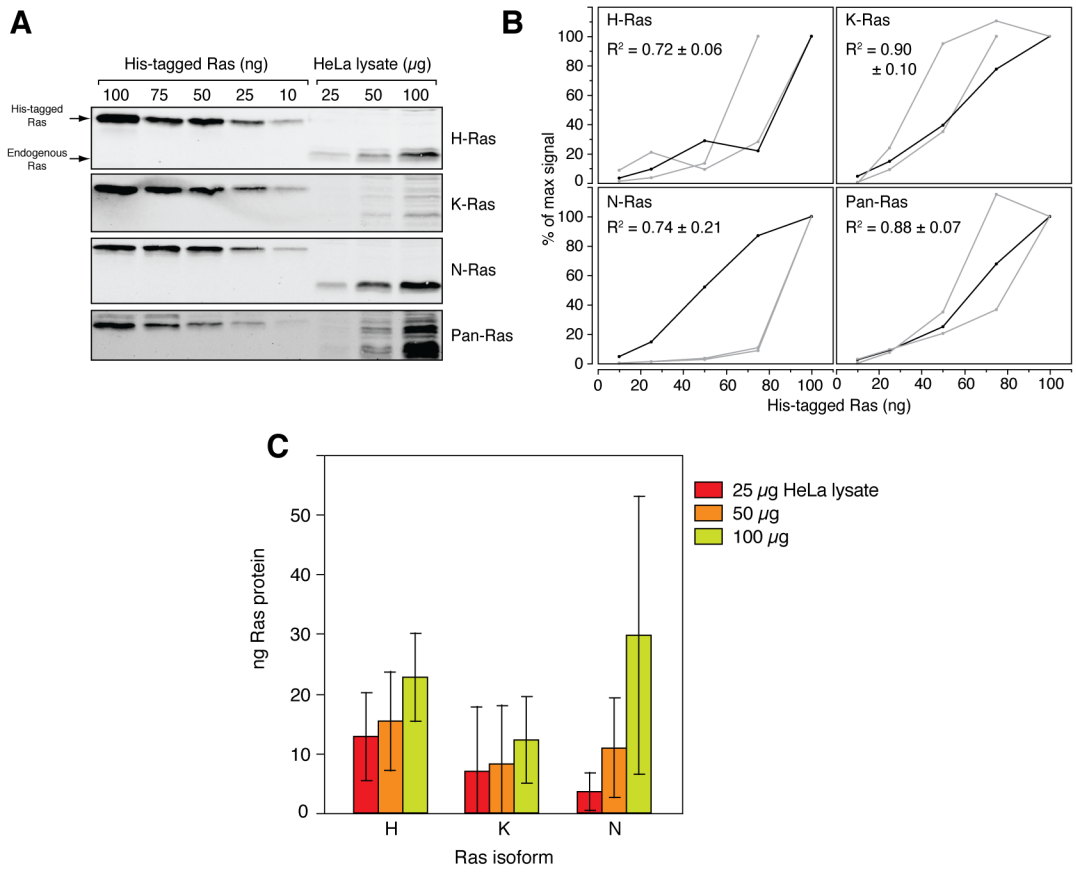


Fig. 3.16. Immunoblot analysis of cellular Ras abundance. (**A**) Immunoblots estimating cellular Ras abundance in HeLa cells. 100 to 10 ng of his-tagged Ras isoforms (1:1:1 mix of H-, K(B)- and N-Ras for pan-Ras) and 25, 50 and 100 μg of HeLa lysate were immunoblotted with isoform-specific and pan-Ras antibodies. (**B**) Response between his-tagged Ras protein abundance and densitometry signal. Black curves correspond to the immunoblots in A. (**C**) Calculated Ras isoform abundance in HeLa cell lysate. Bars represent mean ± SD of 3 biological replicates.

Furthermore, the amount of total Ras calculated using the pan-Ras antibody did not match the individual isoforms summed together, e.g. for 50 μg lysate, the pan-Ras antibody signal indicates that total Ras is 80.59 ± 55.39 ng whereas the summed total is 34.47 ng. Moreover, the quantity of Ras does not double when the amount of HeLa lysate is doubled, from 25 to 50 μg lysate, calculated H-Ras increased 1.2-fold, K-Ras 1.18-fold and N-Ras 3-fold (Fig. 3.16C).

Although immunoblotting generates results with high variability, it does indicate that K-Ras may be the least abundant isoform in HeLa cells, while H- and N-Ras are expressed at similar levels. Nevertheless, this strategy cannot be used to accurately and robustly measure the abundance of Ras isoforms in cell lines or other biological samples.

3.2.2 DETECTION OF PROTEOTYPIC RAS PEPTIDES IN ISOGENIC SW48 COLORECTAL CELL LINES

The present thesis targeted an isogenic panel of SW48 colorectal cell lines, harbouring a range of Ras mutations, for the measurement of cellular Ras abundance. The isogenic cell lines were generated through targeted homologous recombination, which exchanges an endogenous Ras allele for a mutant Ras sequence. In addition to the K-Ras wild-type control cell line (known as the Parental/ wt cell line), the K-Ras mutations present in the isogenic panel represent those observed in over 75% of colon tumours (G12D, G12V and G13D), plus others that are less prevalent (G12C, G12S, G12A and G12R) (Prior et al., 2012; Vaughn et al., 2011). Also included were H-RasG12V and N-RasG12V mutations, which are rare occurrences in colon cancer (< 3%). All mutant cell lines were heterozygous for their respective Ras mutations.

To discover if the selected proteotypic Ras peptides were detectable at endogenous levels in the isogenic SW48 cell line panel, 100 μg of SW48 Parental, K-RasG12D, K-RasG12V and K-RasG13D cell lysates were subject to SDS-PAGE. The Ras containing band was excised, before reduction and alkylation with DTT and IAA, respectively. The sample was then in-gel digested with trypsin (and elastase in a separate digestion), ahead of peptide extraction, chromatographic separation and SRM analysis on the 4000 QTRAP (AB SCIEX), using the settings described in the previous chapter. All peptides were detectable at endogenous levels, bar the K(A)-Ras-specific peptide and the elastase pan-Ras peptide. Despite increasing dwell times for each transition to 400 msec, neither peptide could be detected using the 4000 QTRAP and were therefore excluded from the analysis. The detectable proteotypic Ras peptides are detailed in Fig. 3.17.

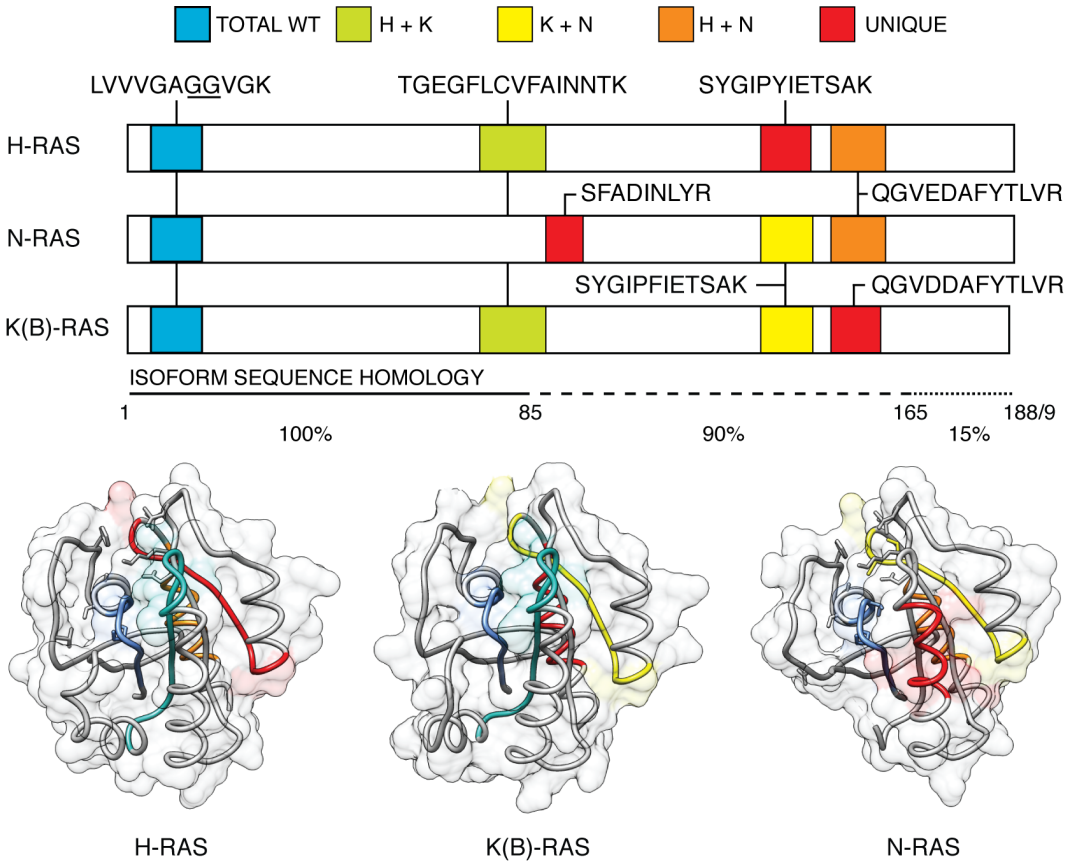


Fig. 3.17. Proteotypic Ras peptides used for quantification of cellular Ras abundance in the isogenic SW48 colorectal cell line panel. Top, the proteotypic, tryptic peptides detected and used for Ras quantification in the SW48 cell line panel. K(A)-Ras was undetectable at endogenous levels in the SW48 cells and was therefore excluded. Bottom, Locations of the proteotypic peptides overlaid on the crystal structures of H-, K(B)- and N-Ras. H-Ras structure from (Milburn et al., 1990), Protein Data Bank (PDB) code: 4q21; K(B)-Ras, PDB code: 3gft; N-Ras, PDB code: 3con, residues 61-71 not visualised. For Ras quantification, 10 ng of isotope-labelled (K8 R10) K(B)-Ras and 5 ng of H- and N-Ras proteins were spiked into 100 μg of cell lysate, before being processed as described in Fig. 3.1.

3.2.3 CELLULAR RAS ABUNDANCE IN ISOGENIC SW48 COLORECTAL CELL LINES

To quantify cellular Ras abundance in the SW48 isogenic panel, cells were grown to ~80% confluency in 10 cm dishes, detached through trypsinisation and resuspended in 10 mL phosphate-buffered saline (PBS), before being counted with a haemocytometer. 6 independent squares on the haemocytometer were counted and only cell counts with < 20% CV were considered for analysis. 9 mL of cells was collected and centrifuged at 200 rcf for 5 min, followed by aspiration of excess PBS. Cell lysis was performed with 200 μL RIPA buffer, supplemented with protease inhibitors, for 20 mins on ice with regular, slight agitation. Cell lysate volume was

measured using a Hamilton syringe, ahead of centrifugation at 13,000 rpm in a bench-top centrifuge for 15 min at 4°C. The cleared lysate was subject to a BCA protein assay, then flash frozen and stored at -80°C until needed. 10 ng of isotope-labelled K(B)-Ras and 5 ng of labelled H- and N-Ras were spiked into 100 μg of each cell lysate, before being subject to SDS-PAGE. A gel band spanning 31-17 kDa was excised and reduced and alkylated, as described above. In-gel tryptic digest was performed overnight and extracted peptides were resuspended 0.1% trifluoroacetic acid (TFA) after drying in a Speed Vac at 60°C. ~1 μg of peptides was separated using RP-HPLC on a NanoAcquity UPLC system before SRM analysis on the 4000 QTRAP triple quadrupole mass spectrometer. Acquired .wiff files were analysed using Skyline (version 2.4).

To ensure that each transition is valid, it was checked that the endogenous and heavy versions of each peptide co-eluted exactly; plus, the order of transitions, with respect to intensity, remained almost identical between endogenous and heavy peptides (Fig. 3.18A). When unlabelled and labelled transitions were integrated, the signal intensity for both versions was within a factor of 3 for all peptides tested (Fig. 3.18B). Together, these observations demonstrate that every transition monitored is genuine and the ratio between light and heavy peptides enables highly accurate quantification. To assess technical reproducibility, 5 independent acquisitions of a single sample were performed, with each peptide showing a variance of ≥ 7.1%.

In the Parental SW48 cell line, per 100 μg lysate, 10.93 ± 0.90 ng of total wild-type Ras was quantified, along with 2.68 ± 0.30 ng H-Ras, 6.00 ± 0.66 K(B)-Ras and 3.38 ± 0.1 ng N-Ras (Fig. 3.19A). Summing the individual isoforms together gives 12.07 ± 0.99 ng, which is 110.4% of the total amount measured by the wild-type Ras peptide. Since the cells were counted and the cell lysate volume measured, it is possible to calculate accurately the number of Ras molecules per cell. In the Parental cell line there is 286,987 ± 33,098 total Ras molecules, with 70,242 ± 6,646 H-Ras, 156,460 ± 18,429 K(B)-Ras and 88,975 ± 3,071 N-Ras molecules. Summing the isoforms together gives 315,676 ± 26,216, which is 110.0% of that calculated using the wild-type peptide (Fig. 3.19A). In the K-RasG12D cell line, 343,782 ± 77,329 wild-type Ras, 116,151 ± 28,943 H-Ras, 297,533 ± 64,544 K(B)-Ras and 155,853 ± 30,385 N-Ras molecules per cell were quantified. Compared to Parental, this represents a 1.65-fold increase in H-Ras, 1.90-fold in K(B)-Ras and 1.75-fold in N-Ras protein copy numbers.

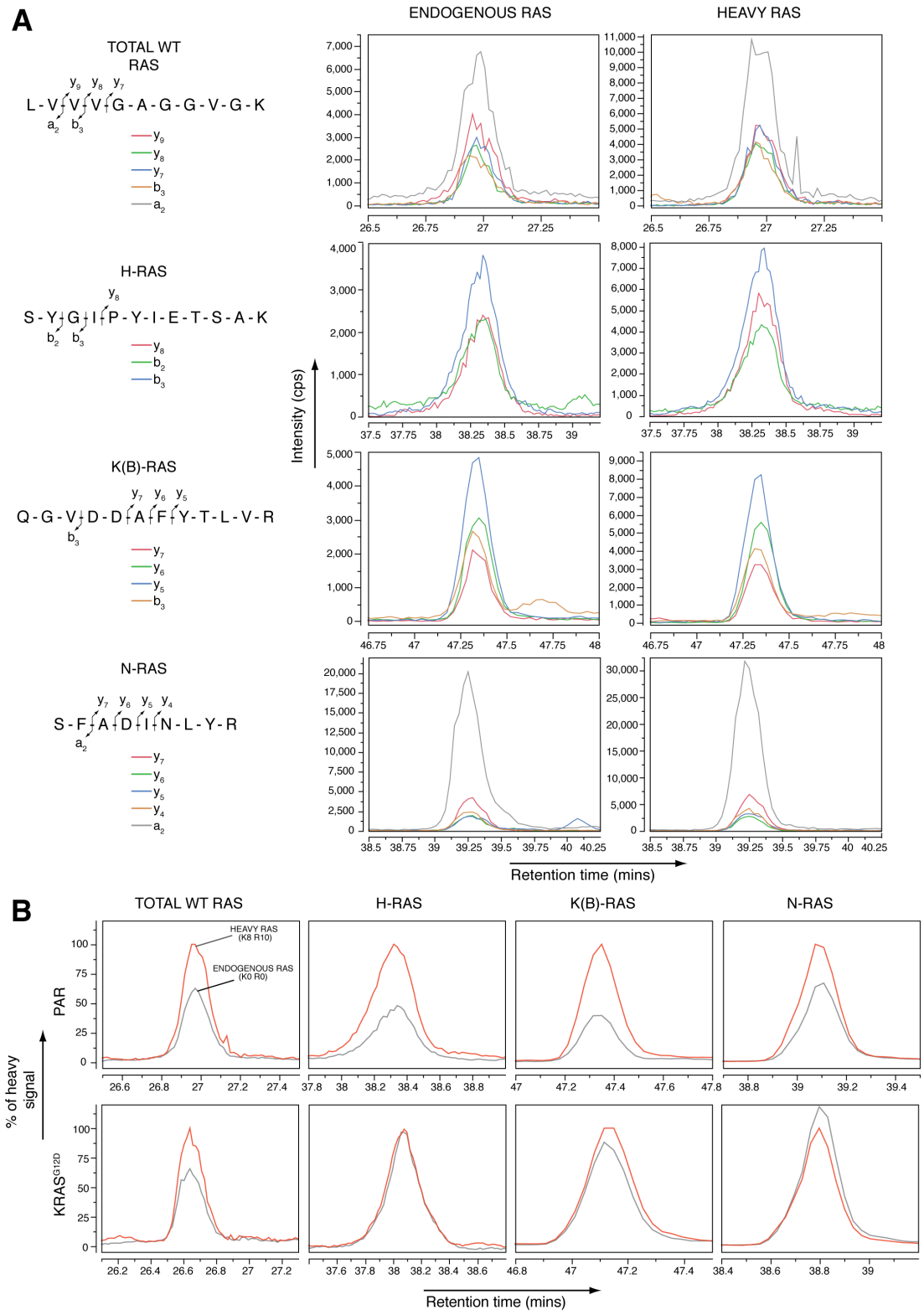


Fig. 3.18. Selected reaction monitoring of proteotypic Ras peptides in SW48 isogenic cell lines. (**A**) Individual transitions of endogenous (K0 R0) and heavy (K8 R10) peptides describing total wild-type, H-, K(B)- and N-Ras abundance in the Parental (homozygous $KRAS^{WT}$) SW48 cell line. (**B**) Integrated transitions (normalised to heavy signal) of the proteotypic Ras peptides in Parental (PAR) and heterozygous KRAS G12D knock-in ($KRAS^{G12D}$) SW48 cells. Presented data representative of 3 biological replicates.

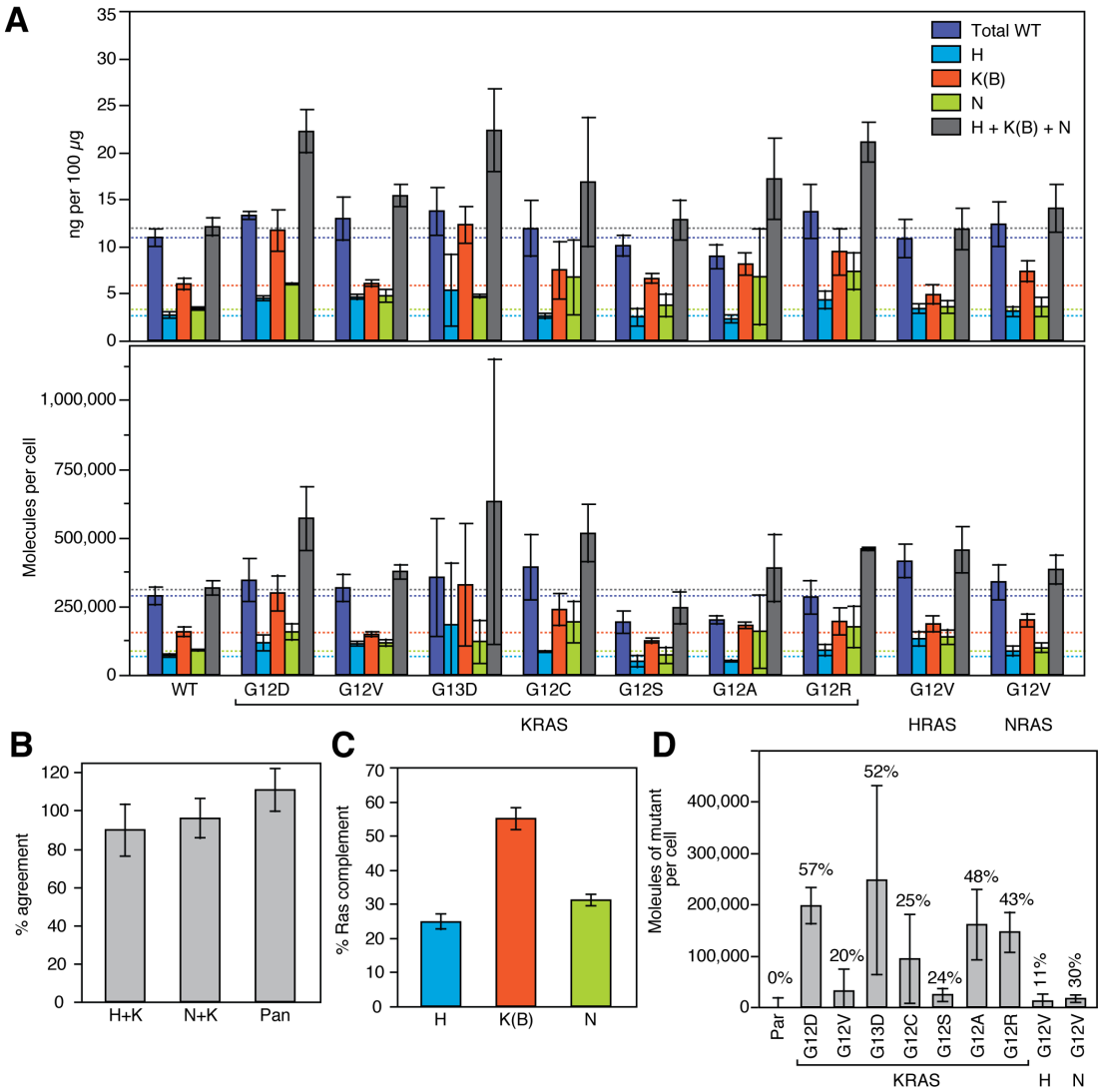


Fig. 3.19. Cellular Ras abundance in SW48 isogenic cell lines. (**A**) Quantify of Ras isoforms in the isogenic colorectal cell lines. Top, ng of Ras per 100 μg of lysate. Bottom, Number of Ras molecules per cell. All mutant cell lines are heterozygous knock-ins. (**B**) Agreement between Ras quantification using peptides shared between isoforms, TGEGFLCVFAINNTK (H and K), SYGIPFIETSAK (K and N) and LVVVGAGGVGK (all wild-type isoforms), with summing the respective isoforms together. $n \geq 20$ for H+K and N+K, n = 3 for Pan as only applicable to the wild-type Ras SW48 cell line. Bars represent mean ± SD. (**C**) Percentage contribution of isoforms to the cellular Ras complement in the wild-type Ras SW48 cell line. (**D**) Number of mutant Ras molecules per cell. Percentages indicate the proportion of the respective isoform that is mutated. Unless stipulated, bars represent mean ± SD of 3 biological replicates. Analysis performed using Skyline (version 2.4).

By adding the individual isoforms together this gives 569,537 ± 115,223 total Ras molecules per K-RasG12D expressing SW48 cell. Subtracting the calculated wild-type, approximately 197,066 ± 61,152 mutant K-RasG12D molecules could be present per cell, meaning ~ 57% of the K-Ras complement is mutated in the K-RasG12D cell line (Fig. 3.19D). The K-RasG13D cell line showed a similar increase in H-, K(B)- and N-Ras numbers as the K-RasG12D cell line, yet had large variance, most likely due to difficulties in counting this cell line. Interestingly, the K-RasG12V cell line had similar Ras levels compared to Parental. Cells harbouring H- and N-RasG12V also demonstrated slight change from the wild-type cell line. The peptides shared between isoforms can be utilised to validate the numbers reported by the peptides unique to isoforms. TGEGFLCVFAINNTK, shared between H- and K-Ras, accounts for 89.83 ± 13.46% of the Ras abundance calculated by H- and K(B)-Ras-specific peptides. While, SYGIPFIETSAK, shared between K- and N-Ras, explains 95.88 ± 10.4% of K(B)- and N-Ras levels (Fig. 3.19B). Although QGVEDAFYTLVR, shared between H- and N-Ras, was monitored, it was unreliably detected during the assay and therefore could not be used. Using the wild-type Parental cell line, the contribution of each isoform to the cellular Ras complement can be calculated: H-Ras 24.70 ± 3.80%, K(B)-Ras 54.97 ± 0.52% and N-Ras 31.10 ± 2.84% (Fig. 3.19C).

3.2.4 DETECTION OF MUTANT-SPECIFIC RAS PEPTIDES

Since the isogenic SW48 colorectal cell line panel express a range of K-Ras mutant proteins, it was tested if mutant-specific peptides could be identified. A G12V-specific mutant peptide, LVVVGAVGVGK, and a G12D/G13D peptide, LVVVGADGVGK, have been detected previously (Ruppen-Canas et al., 2012; Wang et al., 2011). However, the G12D/G13D peptide could not differentiate between codon 12 or 13 aspartate substitutions; plus, peptides unique to other Ras mutants have not yet been discovered. 100 μg of lysate from each SW48 isogenic cell line, harbouring G12D, G12V, G13D, G12A, G12C, G12S and G12R heterozygous K-Ras mutations, was subject to SDS-PAGE and in-gel tryptic digestion as described earlier (LysC digestion for G12R). ~1 μg of peptides was fractionated by RP-HPLC, ionised and subject to MIDAS on a 4000 QTRAP, monitoring for y9, y8 and y7 fragment ions of each mutant peptide.

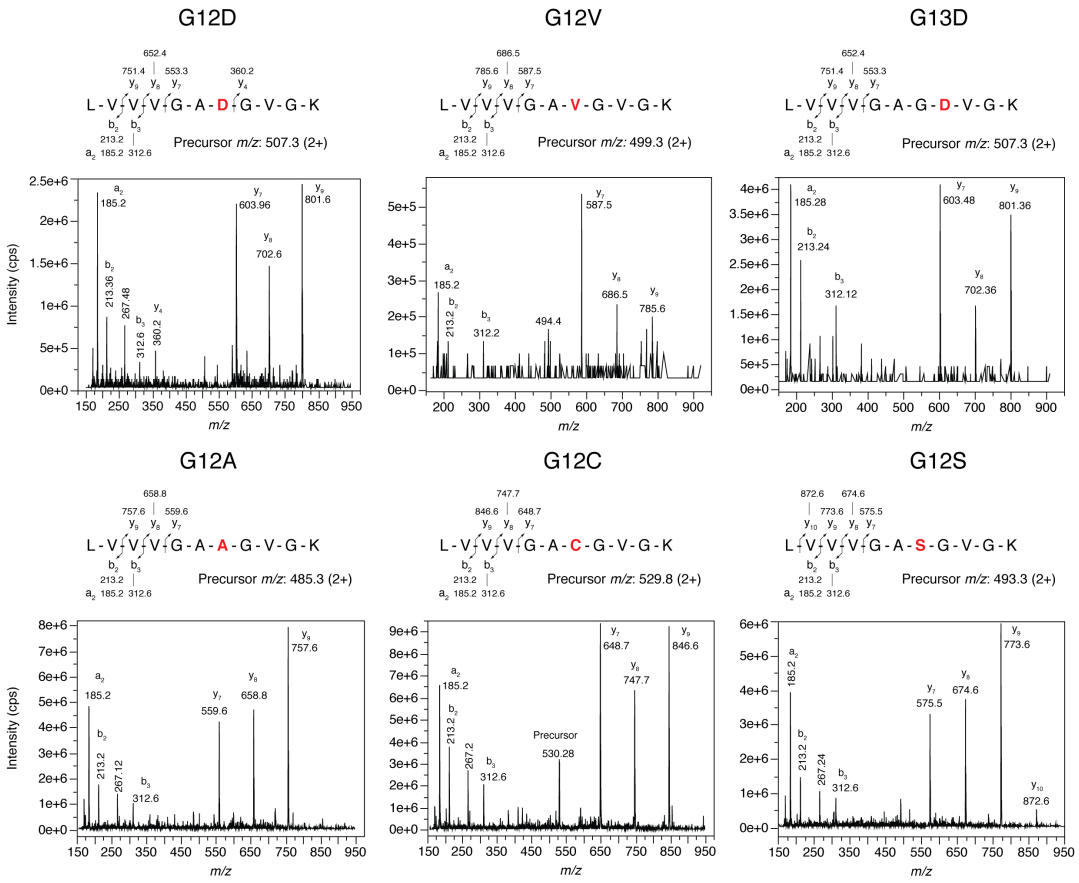


Fig. 3.20. MIDAS detection and MS/MS spectra of mutant-specific Ras peptides. Isogenic SW48 colorectal cell lines, harbouring G12D, G12V, G13D, G12A, G12C and G12S heterozygous K-Ras mutations, were targeted for their respective mutant peptides using tryptic digestion and MIDAS. 100 μg of cell lysate was subject to SDS-PAGE and the region containing endogenous Ras excised, before in-gel tryptic digestion. ~200 ng of peptides were fractionated by RP-HPLC and y9, y8 and y7 ions of each mutant peptide targeted using MIDAS with 200 msec dwell times on a 4000 QTRAP (AB SCIEX). Mutant Ras precursor and product ions were predicted using Skyline (version 2.4). The G12R-specific peptide could not be detected in either doubly or triply charged states. The cysteine residue of the G12C-specific peptide was subject to carbamidomethylation.

Mutant-specific peptides for all mutant proteins were detected, bar the G12R-specific peptide (Fig. 3.20). Moreover, the MS/MS spectrum of the G12D-specific peptide contains a diagnostic y4 ion that differentiates it from the G13D-specific peptide. The G12V-specific peptide was barely detectable with MIDAS, exemplified by its ~10-fold less intensity; plus, it could not be reliably detected using SRM. Detection of these peptides verifies, at the proteomic level, that each cell line expressed its respective K-Ras mutant protein, yet the G12R-specific peptide was undetectable in the K-RasG12R cell line, most likely due to issues associated with the presence of a mid-chain arginine.

3.2.5 TURNOVER ANALYSIS OF RAS WILD-TYPE AND MUTANT PROTEINS

It is unclear if there are differences between wild-type and mutant Ras protein turnover, synthesis and degradation rates. Consequently, SW48 isogenic cell lines harbouring G12V H-, K- and N-Ras mutations, along with G12D and G13D K-Ras mutations, were subject to dynamic SILAC analysis (Fig. 3.21A). To assess turnover, synthesis and degradation rates of the Ras proteins, SW48 cells were grown in 'light' (K0 R0) or 'medium' (K4 R6) media, with the latter pulse labelled with 'heavy' media for various time points. The 'light' and 'medium/ heavy' cells were then mixed at a 1:1 protein ratio and subject to SRM analysis (Fig. 3.21B) (Boisvert et al., 2012). The targeted peptides and selected transitions are detailed in Table 3.3.

To ensure the 'light' and 'medium' cells were appropriately labelled, 50 μg of lysate from each cell line labelled with these isotopes were subject to SDS-PAGE and the Ras-containing band excised and in-gel digested with trypsin. By targeting the total-wild type and K(B)-Ras specific peptides with SRM, labelling of arginine and lysine residues was determined (Fig. 3.22). The 'light' cells had no trace of medium isotopes, while 'medium' cells were >99% labelled on lysine residues and >95% labelled on arginine. 50 μg of 'light' and 'medium/ heavy' cell lysates were mixed and solubilised with sample buffer, prior to SDS-PAGE and in-gel tryptic digest of the Ras-containing gel region. Extracted peptides were subject to the same separation and MS analysis as described earlier. Example SRM traces from the dynamic SILAC experiment are provided in Fig. 3.23. Although 'light', 'medium' and 'heavy' signals of each peptide were reliably detected, the ratios between medium/ light and heavy/ light were highly variable. This meant that, unfortunately, Ras isoform synthesis and degradation rates could not be calculated. However, reliable heavy/ medium ratios were acquired, enabling measurement of Ras isoform turnover (Fig. 3.24). In the Parental cell line, 50% of total wild-type Ras turned over in 15.22 ± 0.91 hours, H-Ras 22.82 ± 2.49, K(B)-Ras 17.49 ± 2.80 and N-Ras 17.46 ± 2.03 hours. The cell lines harbouring K-Ras mutations exhibited slightly increased turnover times for all isoforms. Interestingly, the G13D mutant peptide exhibited a 50% turnover rate of 22.51 ± 1.23 hours, more than a 5-hour increase compared to the wild-type K-Ras in the Parental cell line. The G12D mutant peptide had a lower 50% turnover rate of 17.10 ± 3.29 hours.

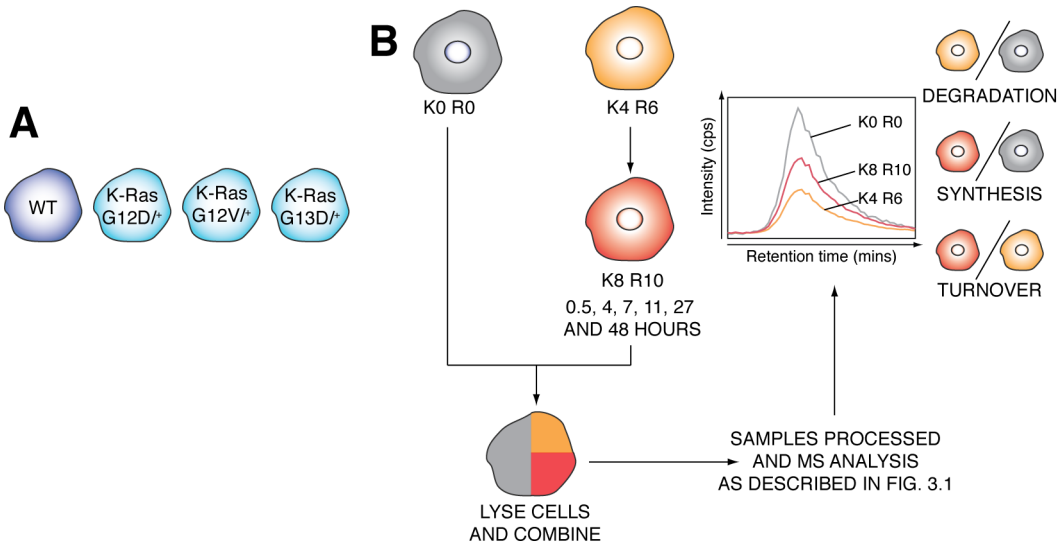


Fig. 3.21. Dynamic SILAC analysis of Ras isoform turnover. (**A**) The SW48 isogenic colorectal cell lines subject to turnover analysis. All mutant cells are heterozygous knock-ins. (**B**) Method employed to determine turnover, synthesis and degradation of Ras isoforms. 'Medium' cells (K4 R6) were exposed to 'heavy' media (K8 R10) for the indicated time periods, before lysis and mixed 1:1 with 'light' cell lysate (K0 R0). Sample processing is described in Fig. 3.1, but no cell counting was performed. Targeted peptides and transitions employed in the selected reaction monitoring analysis are detailed in Table 3.3.

	PEPTIDE	CHARGE	PRECURSOR ION (m/z)			FRAGMENT ION (m/z)			ION	CE (eV)	DWELL TIME (msec)
			K0 R0	K4 R6	K8 R10	K0 R0	K4 R6	K8 R10			
Wild-type Ras	LVVVGAGGVGK	+2	478.3	480.3	482.3	743.4	747.4	751.4	y9	20	100
						644.4	648.4	652.4	y8	22	100
						545.3	549.3	553.3	y7	24	100
Mutant G12D/G13D Ras	LVVVGADGVGK	+2	507.3	509.3	511.3	801.4	805.4	809.4	y9	20	200
	LVVVGAGDVGK					702.4	706.4	710.4	y8	22	200
						603.3	607.3	611.3	y7	24	200
H-Ras	SYGIPYIETSAK	+2	664.8	666.8	668.8	908.5	912.5	916.5	y8	28	200
						308.1	308.1	308.1	b3	34	200
						251.1	251.1	251.1	b2	42	200
K(B)-Ras	QGVDDAFYTLVR	+2	692.4	695.4	697.4	798.5	804.5	808.5	y6	36	200
						651.4	657.4	661.4	y5	36	200
						285.2	285.2	285.2	b3	36	200
N-Ras	SFADINLYR	+2	549.8	552.8	554.8	864.5	870.4	874.5	y7	25	200
						793.4	799.4	803.4	y6	27	200
						565.3	571.3	575.3	y4	25	200

Table 3.3. Targeted peptides and transitions used for Ras turnover analysis. Transitions selected for unlabelled Ras (K0 R0), 'medium' labelled Ras (K4 R6) and 'heavy' Ras (K8 R10). K0, isotopically light lysine; R0, isotopically light arginine; K4, 4,4,5,5-D_4 L-lysine (+4 Da); K6, $^{13}C_6$ L-arginine (+6 Da); K8, $^{13}C_6$ $^{15}N_2$ L-lysine (+8 Da); R10, $^{13}C_6$ $^{15}N_4$ L-arginine (+10 Da). For schematic of the turnover analysis, see Fig. 3.21.

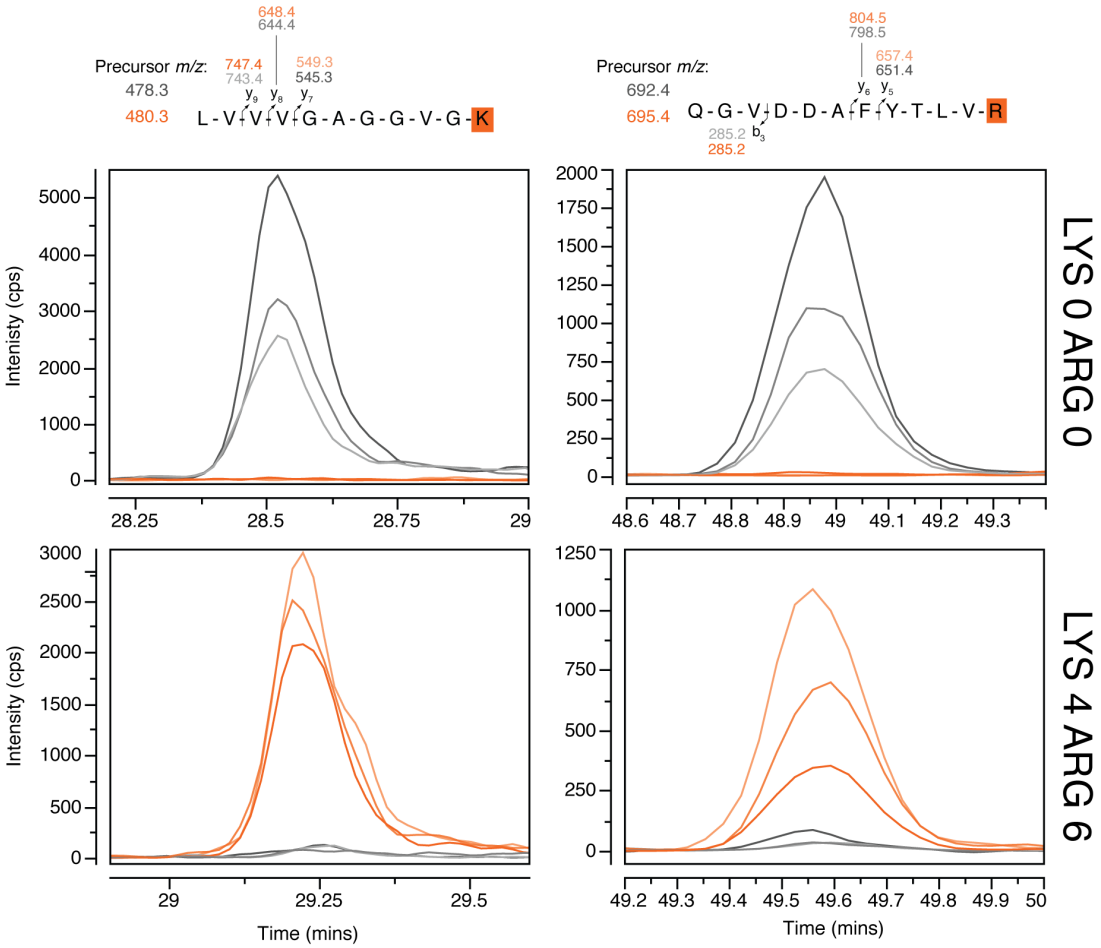


Fig. 3.22. Incorporation of isotope labels into 'light' and 'medium' cells for the Ras turnover analysis. 'Light' cells (K0 R0) had no detectable 'medium' isotope signals, while the 'medium' cells (K4 R6) were > 99% labelled on lysine residues and >95% on arginine residues.

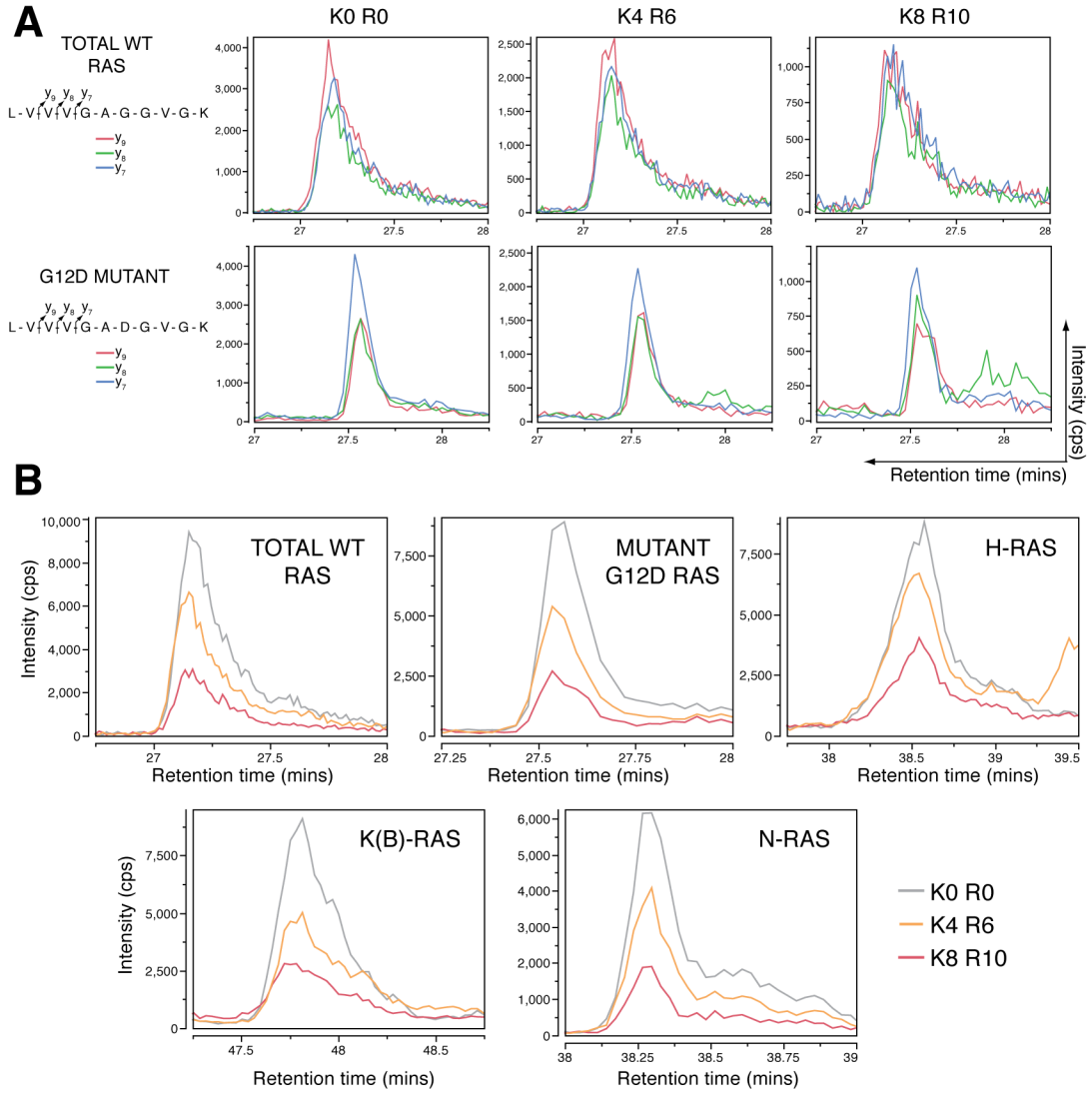


Fig. 3.23. Dynamic SILAC analysis of Ras isoform turnover using selected reaction monitoring. Cells were labelled with 'light' (K0 R0) or 'medium' (K4 R6) isotopes, before the 'medium' cells were pulsed labelled with 'heavy' (K8 R10) media for various time points. Equal protein amounts of 'light' and 'medium/ heavy' cell lysates were mixed, before sample processing as described in Fig. 3.1. (**A**) Individual transitions for wild-type and G12D mutant peptides corresponding to 'light' (K0 R0), 'medium' (K4 R6) and 'heavy' (K8 R10) conditions. (**B**) Example integrated transitions. Taken from 'medium'-labelled SW48 cells, harbouring heterozygous K-RasG12D mutation, pulse labelled with 'heavy' media for 11 hrs.

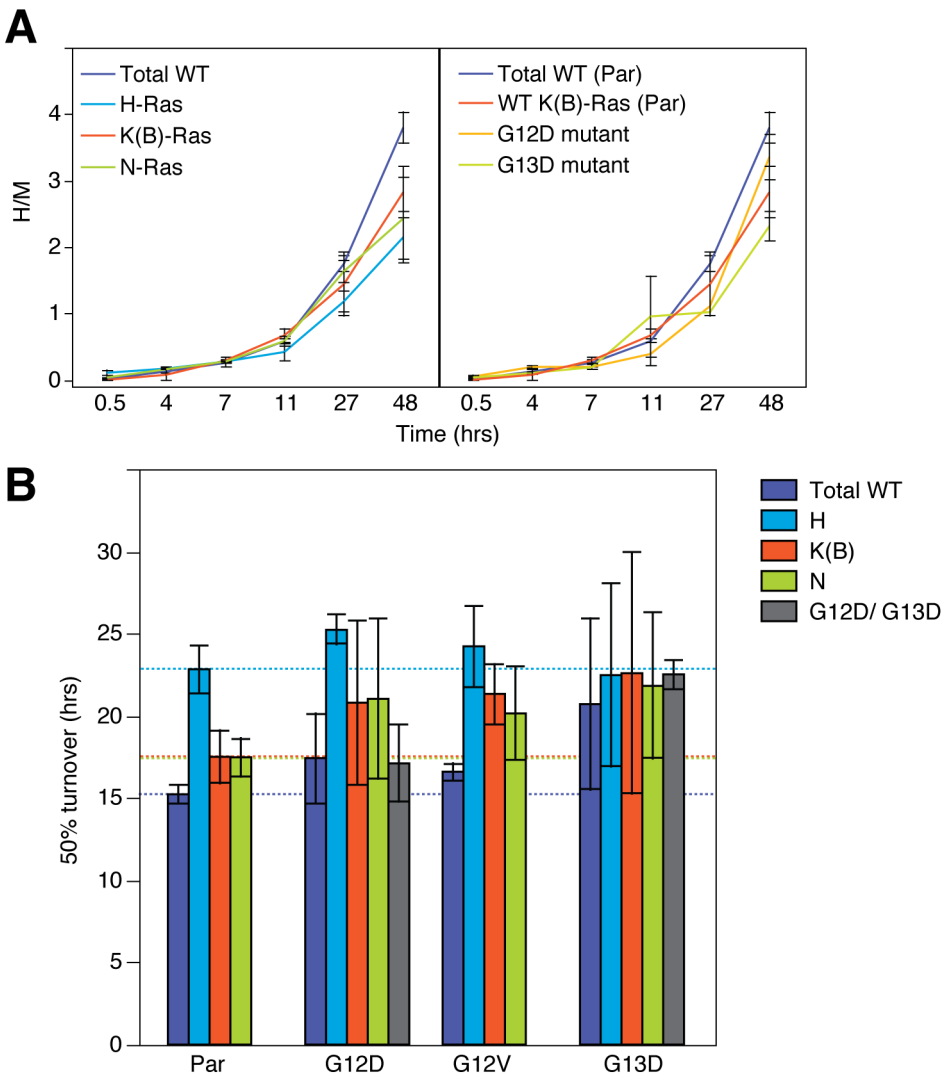


Fig. 3.24. Turnover of Ras isoforms in isogenic SW48 cell lines harbouring various K-Ras mutations. Cells were labelled with 'medium' (K4 R6) media before pulse labelling with 'heavy' (K8 R10) media for various time points. (**A**) Turnover of Ras isoforms and mutant proteins. Left, heavy/ medium ratio (H/M) over the course of the dynamic SILAC experiment in SW48 cells harbouring wild-type Ras (Par). Right, turnover of mutant peptides (G12D and G13D) in cells harbouring these mutations, compared to wild-type K-Ras and total wild-type from the Par cell line. (**B**) 50% turnover of Ras isoforms. 50% turnover was defined as the time point when 'medium' and 'heavy' signals reached a ratio of 1. $n = 3$ for Parental cell line; $n = 2$ for cell lines harbouring K-Ras mutations. Bars indicate mean ± SD.

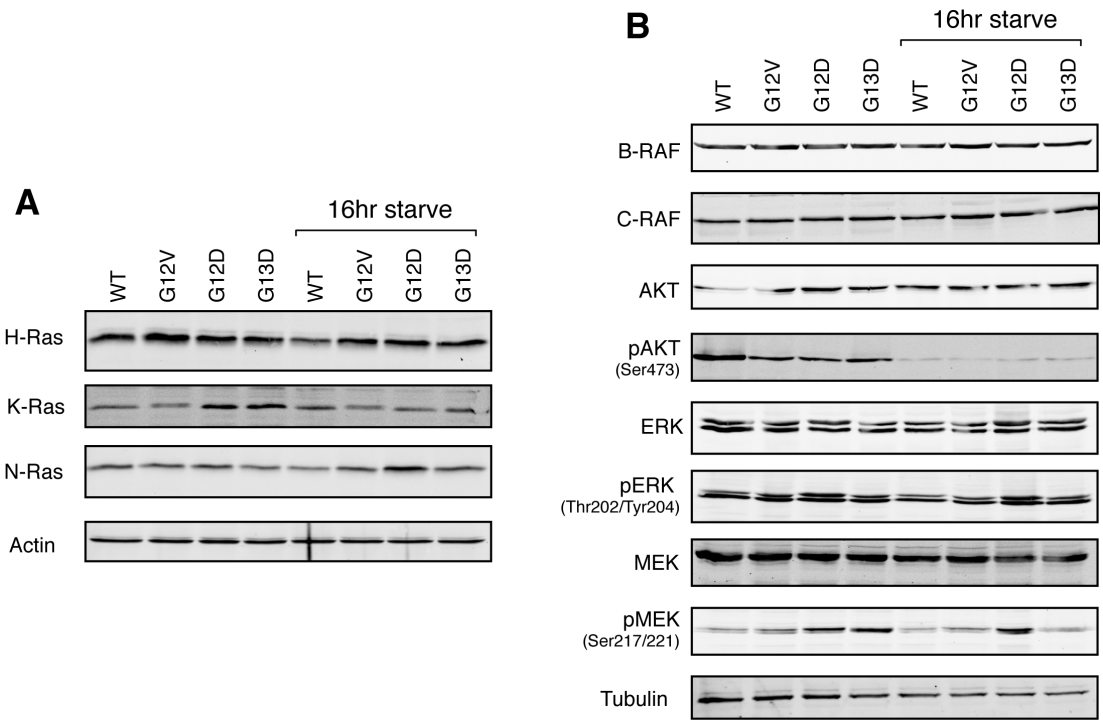


Fig. 3.25. Immunoblot characterisation of colorectal SW48 isogenic cell lines harbouring G12V, G12D and G13D K-Ras mutations. Cells were grown in McCoy's 5A media supplemented with 10% FBS and starved overnight in media with no FBS where indicated (16 hr starve). Cell lysis was performed with RIPA buffer, before 50 μg of SW48 cell lysate was subject to SDS-PAGE and immunoblotting against the indicated proteins. (**A**) Abundance of H-, K- and N-Ras in the isogenic cell line panel. (**B**) Similar activation of AKT and ERK by the different codon Ras mutants, but not MEK.

3.2.6 IMMUNOBLOT CHARACTERISATION OF THE SW48 ISOGENIC CELL LINE PANEL

Isogenic SW48 colorectal cell lines harbouring G12V, G12D and G13D K-Ras mutations were subject to characterisation through immunoblotting. Although immunoblotting was shown to be unreliable when measuring the exact abundance of cellular Ras, it was used to gauge the relative abundance of the Ras isoforms in the SW48 cell line panel (Fig. 3.25A). A slight increase in H-Ras expression is visible from Parental in each of the K-Ras mutant cell lines, while little change is observed in N-Ras abundance. However, there is a clear increase in K-Ras expression in the G12D and G13D K-Ras mutant cell lines. Probing immunoblots for components of canonical Ras signalling pathways revealed a drop in Akt activation in K-Ras mutant cells, while activation of ERK remained constant in all cell lines (Fig. 3.25B). Yet, G12D and G13D mutant K-Ras potently activated MEK.

3.2.7 DISCUSSION

In an initial attempt to quantify cellular Ras abundance, the well-established technique of immunoblotting was utilised. However, inconsistent results were obtained between biological replicates and linear responses between protein abundance and densitometry signal were difficult to obtain, possibly because of issues during transfer or that the antibody cannot discriminate between changes in protein abundance at low protein levels. Consequently, an SRM-based Ras quantification technique was established (as detailed in chapter 1). Although previous attempts have been made to measure cellular Ras abundance with SRM-based techniques, the described methods have disputable accuracy. Wang et al. quantified over 1 million Ras molecules per SW480 cell, by immunoprecipitating Ras proteins from a target lysate, before tryptic digestion and spike in of AQUA peptide standards (Wang et al., 2011). Since their pull-down strategy captured ~25% of Ras protein from samples, and some Ras remained attached to the beads following elution, a correction factor had to be applied to the acquired data. The establishment of this correction factor is confusing, as it was performed by spiking isotopically light Ras protein at different concentrations into an unlabelled cell lysate, which inexplicably already had its Ras abundance measured. Moreover, no useful information is given regarding how the cells were counted in the study. Furthermore, quantification was enabled by AQUA, which assumes complete digestion of the endogenous material. Ruppen-Canas et al. employed the same technique as the Wang study to measure Ras abundance in various tumour samples, but quantified Ras at ~100-fold lower levels (Ruppen-Canas et al., 2012). Consequently, two independent laboratories have utilised the same technique, but obtained two very different results despite analysing similar samples. Halvey et al. adopted a different strategy by subjecting a target lysate to SDS-PAGE before in-gel tryptic digestion, extraction of peptides and spike in of AQUA peptide standards (Halvey et al., 2012). This approach does not account for different efficiencies and fluctuations in peptide recovery from gel pieces, nor does it account for incomplete digestion of the endogenous material. Global, label-free approaches employing the intensity based absolute quantification (iBAQ) algorithm also provide an insight into cellular Ras abundance. Schwanhausser et al. and Nagaraj et al. measured the abundance of thousands of proteins in NIH3T3 and HeLa cell, respectively (Nagaraj et al., 2011; Schwanhausser et al., 2011). Although the iBAQ technique enables the quantification of thousands of proteins, it can only provide approximate protein copy

numbers. This is because the number of theoretically observable peptides, a parameter used by the algorithm to quantify total protein copy numbers, is predicted computationally and may not be correct for many proteins. In addition, each peptide has its own unique MS response, meaning that most peptides, even from the same protein or present at the same concentration, will be detected at different intensities by the mass spectrometer. Consequently, peptide physicochemical properties can undermine the accuracy of the iBAQ technique, if only one or two peptides have been detected for a protein. The iBAQ approach is most useful to describe the abundance of protein groups as a whole (personal correspondence, Matthias Selbach, Max Delbrück Center for Molecular Medicine, Germany). Here Ras abundance was measured using the PSAQ strategy, which obviates the need for a correction factor and resolves problems associated with incomplete digestion and fluctuations/ differences in peptide recovery from gel pieces; plus, all measured endogenous Ras peptides are compared to a chemically identical, isotope-labelled counterpart. Consequently, the presented data represents the most accurate measurement of cellular Ras isoform abundance to date.

In the Parental SW48 cell line, 10.93 ng of total Ras per 100 μg of lysate was quantified, which is somewhat different to the Halvey study that quantified 1-7 ng of total Ras in 100 μg lysate in all their samples, while the Wang study detected over 25 ng of Ras per 100 μg sample in their cell lines, with the SW480 cell line containing more than 100 ng of Ras. However, it is unclear how Wang and colleagues calculated that there are over 1 million wild-type Ras molecules per SW480 cell. Wang et al. report that an SW480 cell has 155% more wild-type Ras than the Parental SW48 cell lines quantified in this thesis, but obtained 4 mg of protein from the lysis of 5.6 million SW480 cells, whereas the present thesis obtained ~1 mg of protein from ~10 million SW48 cells. Such a protein yield reported by the Wang study is a very impressive one, but since they do not stipulate how their cells were counted, an element of doubt remains over its accuracy and over their quantification technique in general.

The present thesis calculates that there are 287,000 Ras molecules per Parental SW48 cell and assuming a cell's diameter is ~20 μm (diameter determined using an electron microscope grid, performed by Prof. Ian Prior) and that ~50% of Ras protein is associated with the plasma membrane (Laude and Prior, 2008), it can be

calculated that per μm^2 of plasma membrane there are approximately 114 bound Ras molecules. As the volume of a Parental SW48 cell is ~2 pL (measured using a Scepter™ cell counter, Millipore) the cellular molarity of Ras is 253 mM. The cellular Ras complement was also determined, with K(B)-Ras the most abundant isoform (55%), followed by N-Ras (31%) then H-Ras (25%). Since these final values are subject to the summed technical and biological variability of 4 independent peptide measurements from 3 biological replicates, it is not unexpected or concerning that they do not add up to 100% exactly (actually add up to 111%). K(A)-Ras was undetectable in the SW48 cell line, which could be due to a little to no mRNA expression for this minor K-Ras isoform (Plowman et al., 2006).

Compared to the Parental cell line, cells harbouring G12D and G13D K-Ras mutations express increased levels of H- and N-Ras, and nearly double the amount of K(B)-Ras, which was also observed using immunoblotting. This is interesting since G12D and G13D are first and third most common K-Ras mutation in CRC, respectively (Vaughn et al., 2011). By summing the individual isoforms together, the K-RasG12D mutant cell line contains 1.8-fold more Ras compared to Parental, with as many as 197,000 mutant Ras proteins per cell, meaning around 57% of the K-Ras protein expressed is mutant. G12A and G12R K-Ras mutations also show an increase in K-Ras protein expression, but not at the same levels as G12D and G13D. In contrast, the second most common K-Ras mutation in CRC, G12V, exhibited slight change in Ras isoform abundance compared to Parental, indicating little K-RasG12V protein expression. Yet the mutant protein is present in these cells, since a G12V-specific mutant peptide was detected using MIDAS, with 31,000 mutant K-RasG12V proteins per cell, accounting for 20% of K-Ras complement. Similarly, neither H- nor N-RasG12V mutations induced significant change to Ras isoform expression compared to Parental. These data suggest that mutant Ras protein expression is associated with its transforming activity. Early Ras research showed that different H-Ras mutations exhibited a range of transforming activities, with G12V usually having the greatest potency (Der et al., 1986; Seeburg et al., 1984; Sloan et al., 1990). K-RasG12V reduced activation of the PI3K pathway to a similar extent as G12D and G13D mutations, despite having 15-fold lower expression, but could not activate MEK like these mutations. This, it could be that Ras mutations exert their different biological effects, not only through slight structural differences or differently impaired GTPase activities, but also through

differential protein expression. Furthermore, since a Ras mutation induces stress on the cell, a K-RasG12V mutation could be such a prolific oncogenic signalling molecule that the cell cannot tolerate its increased expression and actively limits its cellular abundance. While G12D and G13D mutations could be less transforming than G12V, as eluded to previously (Cespedes et al., 2006), enabling the cell to tolerate the presence of hundreds of thousands of these mutant molecules.

Each cell line harbouring a heterozygous Ras mutation expressed similar levels of wild-type Ras protein as Parental. This could indicate that the mutant Ras protein is simply added on top of the usual cellular Ras complement in the SW48 cell line panel. It also shows that despite a range of Ras mutations being present, the cell maintains a minimum level of wild-type Ras. Mutant and wild-type Ras have been shown to interact with each other, with mutant Ras proteins able to activate endothelial nitric oxide synthase or SOS which in turn activates wild-type Ras (Jeng et al., 2012; Lim et al., 2008); plus, wild-type H- and N-Ras can help promote K-Ras-driven tumourigenesis (Bentley et al., 2013; Grabocka et al., 2014; Young et al., 2013). Yet, wild-type Ras proteins have also been reported to act as tumour suppressors, with wild-type H-Ras alleles deleted in some mouse cancers (Bremner and Balmain, 1990; Keller et al., 2007b; To et al., 2013). Thus, the cellular wild-type-to-mutant Ras protein ratio is likely to be context dependent. In CRC, as the colorectal SW48 isogenic panel maintained a minimum amount of wild-type Ras, with increases in wild-type H- and N-Ras in G12D and G13D K-Ras mutant cell lines, wild-type Ras isoforms may contribute to colorectal carcinogenesis in the presence of certain K-Ras mutations.

To determine if Ras isoforms exhibit different turn over rates in the presence of different K-Ras mutations, isogenic SW48 cells harbouring wild-type (Parental), or those harbouring G12D, G12V and G13D K-Ras mutations, were subject to dynamic SILAC analysis. The initial aim was to determine synthesis, degradation and turnover rates, but as the light-to-medium and light-to-heavy ratios were erratic in most analyses, it prevented synthesis and degradation calculations from being performed. The reason for inconsistent ratios is unclear, as all cells were seeded in the same manner and treated with the same media, although containing different isotope labels. However, the medium-to-heavy ratio was reliable throughout the analyses, which provided an insight into how the Ras protein isoforms turnover. In

the Parental cell line, H-Ras had the longest 50% turnover rate, with K(B)- and N-Ras showing similar times. This shows for the first time that Ras isoforms do not share the same turnover rate or half-life. The measured 50% turnovers at a glance seem to be shorter than the expected 20 hour half-life of Ras (Ulsh and Shih, 1984), since the calculations have not considered cell growth. Interestingly, the turnover rates of the Ras isoforms increase in the presence of a K-Ras mutation, but the variability also increases. However, the G13D-mutant peptide displayed a 5-hour longer turnover rate compared to the wild-type K(B)-Ras peptide in Parental cells. Although this observation needs to be verified, it provides the first indication that certain mutant Ras proteins may turnover differently within the cell. This could also provide a mechanism through which mutant Ras proteins can increase their expression levels.

3.3 PROTEOMIC AND PHOSPHO-PROTEOMIC PROFILING OF ISOGENIC COLORECTAL CELLS HARBOURING DISTINCT ACTIVATING K-RAS MUTATIONS

Mammalian cells express three isoforms of the highly homologous p21 Ras family of monomeric GTPases: H-, K- and N-Ras. These proteins act as molecular switches and timers, which are predominantly localised to the plasma membrane where they are activated by cell surface receptors, e.g. epidermal growth factor receptor (EGFR), to regulate cell proliferation, survival and apoptosis (Karnoub and Weinberg, 2008). Mutation of Ras at conserved codons 12, 13 or 61 results in constitutive activation of the molecule, through impaired intrinsic GTPase activity and resistance to GTPase activating protein (GAP) activity (Prior et al., 2012). Although each codon mutation is activating, the position of the mutation has functional and clinical relevance.

Colorectal cancer (CRC) is the third most common cancer worldwide, with an overall 5-year survival rate of 11% for patients with metastatic disease (Siegel et al., 2012). Over a third of CRC tumours harbour a K-Ras mutation (Prior et al., 2012), while the majority of these mutations affect codon 12 (78%), a significant fraction are at codon 13 (22%) (Vaughn et al., 2011). Early Ras research indicated that different amino acid substitutions at codons 12, 13 and 61 exhibited a range of transforming activities (Der et al., 1986; Seeburg et al., 1984; Sloan et al., 1990). More recent work has shown that K-RasG12D and G13D mutations induce different soft tissue sarcoma types and metabolic phenotypes (Guerrero et al., 2002; Vizan et al., 2005), though K-RasG12D more potently transforms cells and makes them more resistant to apoptosis (Guerrero et al., 2000). Yet, CRC patients harbouring K-RasG13D mutant tumours tend to display a worst overall survival rate than patients with other K-Ras mutations (Bazan et al., 2002; Modest et al., 2011; Samowitz et al., 2000; Tejpar et al., 2012). Furthermore, CRC patients with codon 12 and 13 K-Ras mutations display differential responsiveness to anti-EGFR therapy (De Roock et al., 2010; Tejpar et al., 2012). Together, these data suggest that different activating K-Ras mutations are unequal in their signalling outputs. The mechanisms through which this occurs are unclear, but it may be attributable to the slight structural differences between the mutant Ras proteins (Franken et al., 1993; Muraoka et al., 2012), the differentially impaired GTPase activity of the Ras mutants (John et al., 1988; Trahey and McCormick, 1987) or the shift in equilibrium of the GTP-bound

Ras conformational states (Geyer et al., 1996; Spoerner et al., 2001; Ye et al., 2005).

Previous studies have applied various 'omic' approaches to identify mutant K-Ras specific signatures (Guha et al., 2008; Singh et al., 2012; Sudhir et al., 2011; Sweet-Cordero et al., 2005; Tchernitsa et al., 2004; Vartanian et al., 2013). However, some of these studies utilise different cell lines to compare oncogenic Ras signalling, which, due to the variability in genetic background, makes direct attribution of the result to the presence of oncogenic Ras difficult. One strategy to overcome this is use of isogenic cell lines. To date, almost all isogenic models either stably over-express oncogenic Ras, following random insertion of the mutant gene into the genome of an isogenic cell line, or genetically ablate a wild-type or mutant allele.

Here we utilise isogenic SW48 colorectal cell lines, in which targeted homologous recombination has exchanged (knocked-in) an endogenous K-Ras allele for a mutant K-Ras sequence. Since the mutant K-Ras gene is under control of the endogenous K-Ras promoter, this isogenic model enables mutant K-Ras protein to be expressed at physiologically relevant levels. The G12D, G12V and G13D K-Ras mutations present in the isogenic panel are the three most observed K-Ras mutations in CRC, representing over 75% of all CRC cases that harbour a K-Ras mutation (Prior et al., 2012; Vaughn et al., 2011). These isogenic cell lines were subject to quantitative proteomic and phospho-proteomic MS-based analyses, to determine signalling signatures associated with each K-Ras mutation.

3.3.1 K-RAS MUTATION-SPECIFIC NETWORK RESPONSES

Isogenic SW48 colorectal cell lines harbouring homozygous wild-type K-Ras or heterozygous G12D, G12V or G13D K-Ras mutations were used to investigate the effects of different amino-acid substitution (G12D vs. G12V) or codon mutation (G12D vs. G13D). Stable isotope labelling of amino acids in cell culture (SILAC) (Ong et al., 2002) was employed in a triplex conformation to selectively label each of the isogenic cell lines, such that a minimum of 3 biological replicates comparing each mutant cell line with the Parental control cell line was performed (Fig. 3.26). Following SILAC labelling, cell lysates were mixed at a 1:1:1 protein ratio before being subject to SDS-PAGE or TiO_2-based phospho-peptide enrichment strategies.

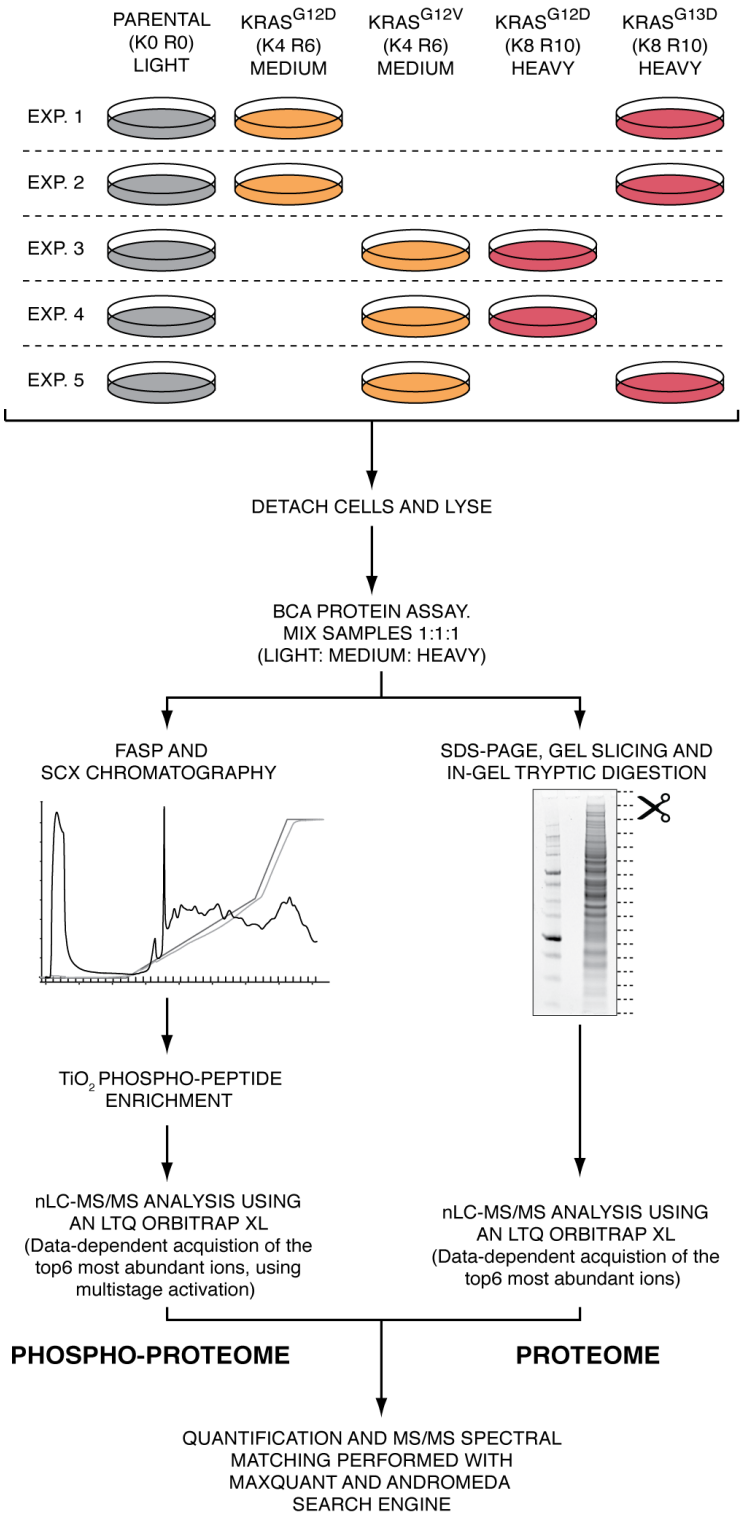


Fig. 3.26. Experimental workflow employed to profile isogenic SW48 colorectal cell lines, harbouring distinct activating K-Ras mutations, at proteomic and phospho-proteomic levels. The isogenic SW48 cell lines were subject to targeted homologous recombination of an endogenous K-Ras allele to knock-in a wild-type (Parental) or a mutant sequence encoding for K-Ras G12V, G12D or G13D. The configuration is such that at least n = 3 for each mutant cell line versus control (n = 4 for K-RasG12D vs. Parental). FASP, filter-aided sample preparation (Wisniewski et al., 2009); SCX, strong-cation exchange. Experiments performed in collaboration with Dr. Dean Hammond (University of Liverpool).

Gel slices and phosphopeptide-enriched fractions were subject to LC-MS/MS analysis using an LTQ Orbitrap XL (Thermo Fisher), in data-dependent acquisition mode sequencing the top 6 most abundant proteins. Phosphopeptide fractions were subject to multistage activation, activating 49.0, 69.3 and 98 Th from the precursor ion mass. Quantification and spectral matching was performed with MaxQuant and the Andromeda search engine (Cox and Mann, 2008; Cox et al., 2011). To enable consistent and predictable cell growth, the isogenic SW48 cells were grown in McCoy's 5A SILAC media. Due to the presence of bacto-peptone, labelling arginine residues was 69.43 ± 0.54% and lysine 89.12 ± 0.13%. Yet, this labelling discrepancy is accounted for during the 'first search' performed by MaxQuant (personal correspondence, Jürgen Cox, Max-Planck-Institut für Biochemie, Martinsried, Germany).

Across all experiments, responses were measured from 2,359 unique proteins in the proteome data and 3,971 unique phosphopeptides from TiO_2 enrichments. 65% of proteins and 35% of phosphopeptides were identified at least twice across biological replicates (Fig. 3.27A). 3,727 phosphosites could be assigned to a specific amino acid (localisation probability ≥ 0.75, score differential > 5; class 1 phosphosite), which comprised of 3,030 pSer, 632 pThr and 65 pTyr sites mapped to 1,288 proteins.

To determine reproducibility between biological replicates and inter-mutation response variability, cross-correlation analysis between all experiments was performed (Fig. 3.28A). Hierarchical clustering indicates good experimental reproducibility between biological replicates, which have at least a 5-fold higher coefficient of determination (R^2) than between any other inter-isogenic cell type correlations. These consistent responses enabled observation of K-Ras mutation-specific signatures. The cross-correlation analyses, along with principle component analysis (Fig. 3.28B), indicates that, whilst proteome and phospho-proteome outputs are similar between G12V and G12D K-Ras, there is divergence between codon 12 and 13 mutations. The major difference between G12V and G12D mutations seem to be the extremity of their responses (Fig. 3.27C and D; Fig. 3.28B). In contrast, G13D mutant cells show prevalent protein and phosphopeptide up-regulation, compared to the codon 12 mutations (Fig. 3.27B and C).

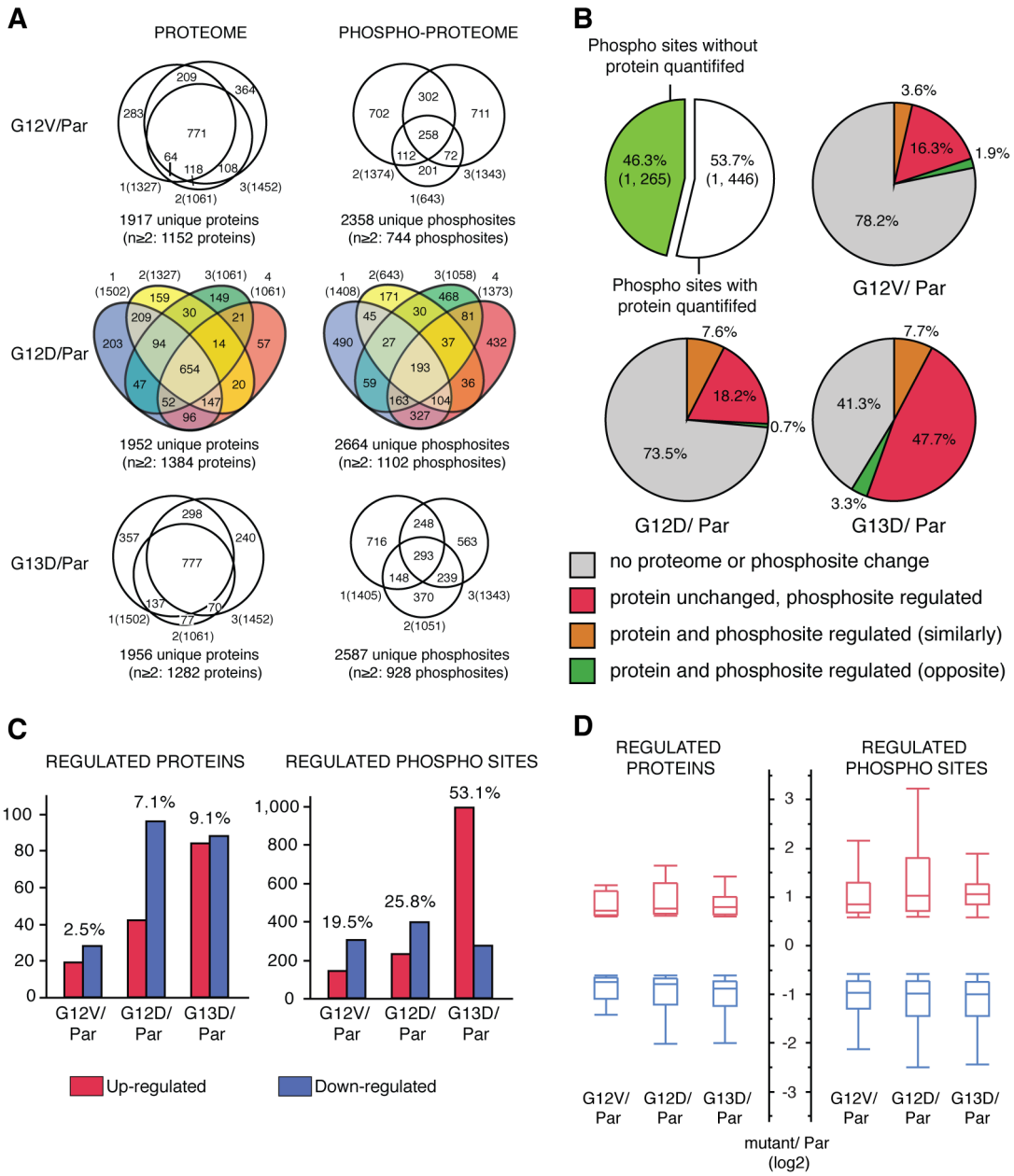


Fig. 3.27. Overview of acquired SILAC data sets. (**A**) Overlap of identified proteins and phosphopeptides between biological replicates. 65% of proteins and 35% of phosphopeptides were observed in more than 1 experiment. (**B**) Overlap and co-responsiveness between proteome and phosphoproteome data sets. (**C**) Responsive subsets within each K-Ras mutant cell line. Percentages indicate the proportion of proteins/ phospho sites significantly changed compared to the total proteome or phospho-proteome. (**D**) Range of regulated protein and phospho site values. Proteins and phospho sites were considered significantly changed/ regulated according to ≥ 1.5-fold change from Parental.

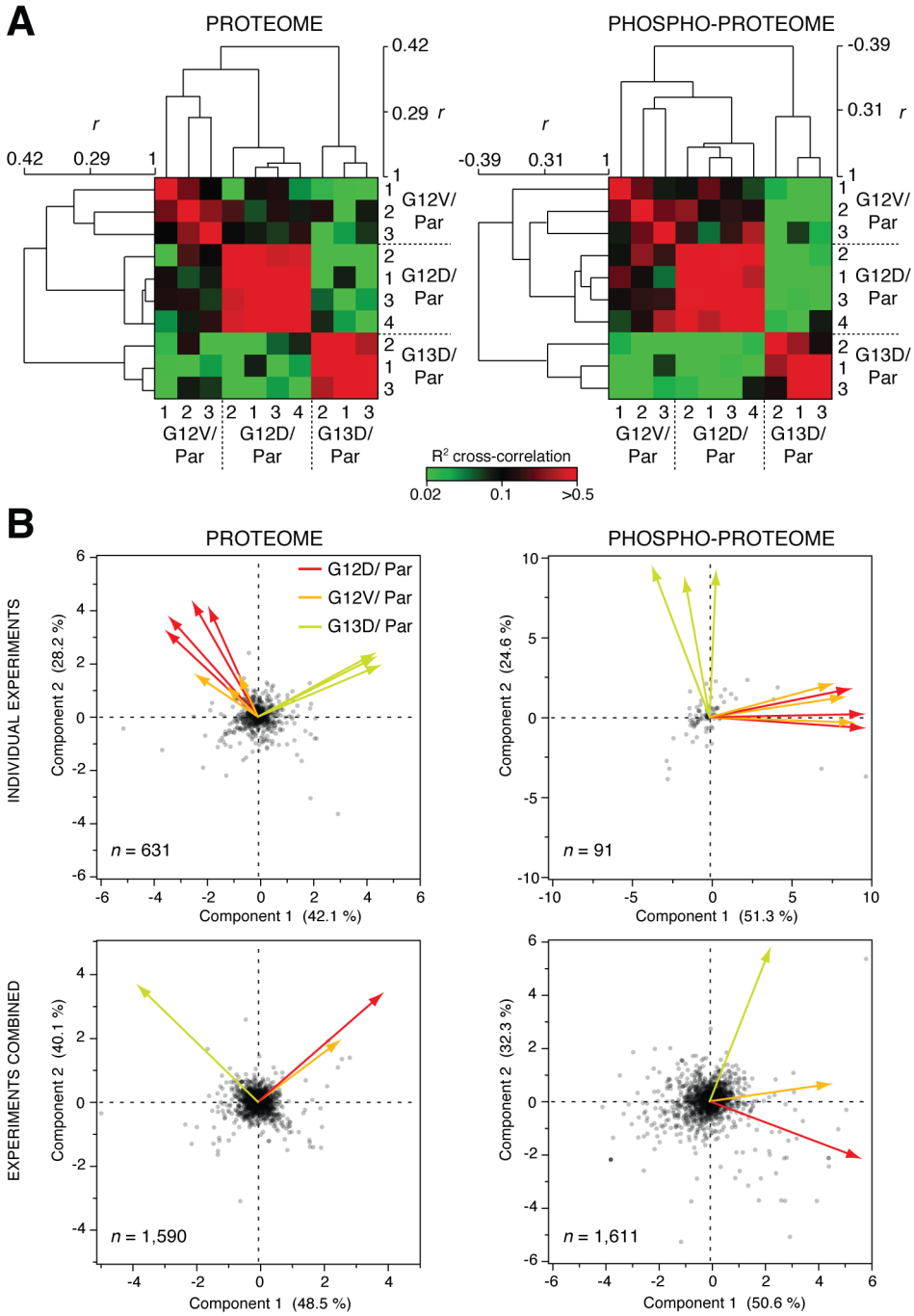


Fig. 3.28. K-Ras mutation-specific proteome and phospho-proteome network responses. (**A**) Cross-correlation analyses between each experiment performed. A consistent correlation between biological replicates is apparent, while cells harbouring codon 12 mutations share greater correlation compared to cells with K-RasG13D, most evident in the phospho-proteome. Values for peptide/ phospho-peptides shared between each experiment were used in the analysis. Pearson correlation (*r*) indicates linkage strength between experiments. (**B**) Principle component analysis on covariances of individual experiments (top) and following combination of the data (bottom). Codon 12 mutations exhibit similar projections at both proteome and phospho-proteome levels, unlike K-RasG13D.

3.3.2 DIFFERENTIAL RESPONSES BETWEEN CODON 12 AND 13 MUTANT K-RAS

Utilising the GProX proteome analysis software (Rigbolt et al., 2011b), proteins and phosphopeptides were clustered according to their mutation-specific responses versus Parental (Fig. 3.29 and Fig. 3.30). Six clusters were generated for each data set, with clusters 5 and 6 being of particular interest since they represent codon 12 versus codon 13-specific responses. In the proteome data set, 115 proteins were isolated in clusters 5 and 6. Gene Ontology (GO) analysis (performed using the DAVID functional annotation tool 6.7) revealed that mitochondrial proteins involved in oxidative phosphorylation were enriched in cluster 5 (Huang da et al., 2009a; Huang da et al., 2009b; Iwata et al., 1998), i.e. proteins up-regulated by codon 12 mutant K-Ras, but not K-RasG13D (Fig. 3.29B). In cluster 6, proteins involved in carbohydrate biosynthesis and cellular homeostasis were enriched.

In the phospho-proteome data set, 274 phosphopeptides from 198 proteins were contained in clusters 5 and 6 (Fig. 3.30A). 56 phosphopeptides from these clusters are associated with 27 cell adhesion or cytoskeleton proteins, with microtubule-associated protein 1B (MAP1B) and tight-junction protein ZO-2 (TJP2) represented by multiple phosphopeptides (Fig. 3.30B and C). Other notable proteins with phosphosites present in clusters 5 and 6 include the HGF-receptor MET T995 and Caveolin-1 S37, which both exhibit a >10-fold increased phosphorylation status compared to codon 13, along with BRAF S729 that exhibits enhanced phosphorylation in codon 13 mutant cells versus codon 12.

To examine complete patterns of phosphopeptide responses, proteins represented by multiple phosphopeptides (≥ 3 in entire data set) from clusters 5 and 6 were curated and compared with available proteome data (Fig. 3.31). Comparison of protein abundance with phosphopeptide signal enables determination if a phosphosite displays increased phosphorylation or that a change protein expression is responsible for the observed alteration in signal. Changes in protein abundance tended to be a minor influence on phosphopeptide ratios, but interestingly protein abundance and phosphorylation typically trended in the same direction.

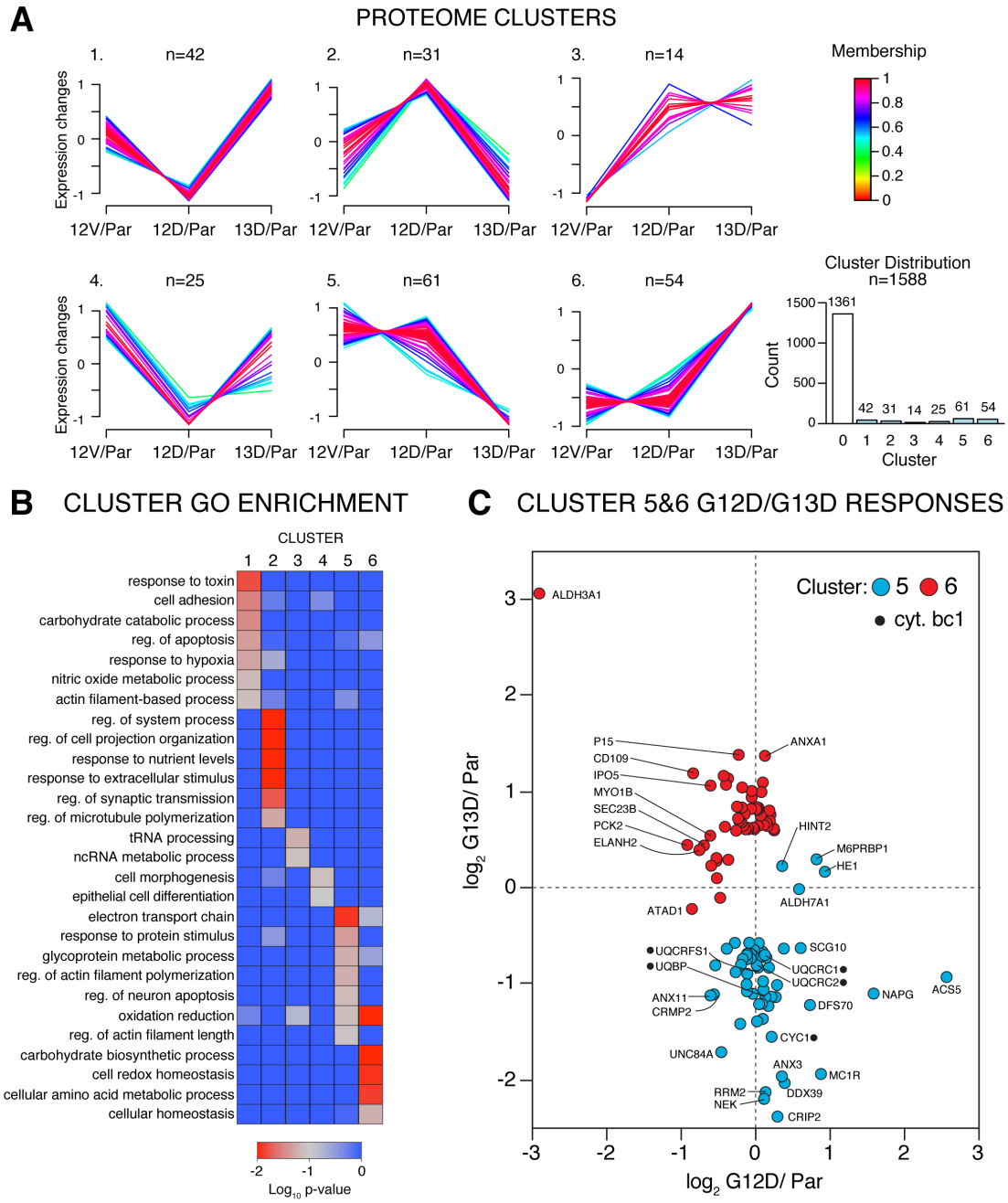


Fig. 3.29. Proteins displaying differential codon 12 and 13 mutant K-Ras responses. (**A**) Clustering of proteins based on response to the different K-Ras mutations. Mutant/ Par ratios were subject to clustering with the Fuzzy c means algorithm using GProX. Clusters 5 and 6 contain proteins that are likely to be signatures of codon 12 and 13-specific signalling in colorectal cancer. (**B**) Gene Ontology (GO) analysis of each cluster performed using the DAVID functional annotation tool 6.7. (**C**) Scatterplot of clusters 5 and 6 with proteins involved in the cytochrome bc1 complex (cyt. bc1) highlighted. GProX cluster analysis performed with the help of Dr. Dean Hammond (University of Liverpool).

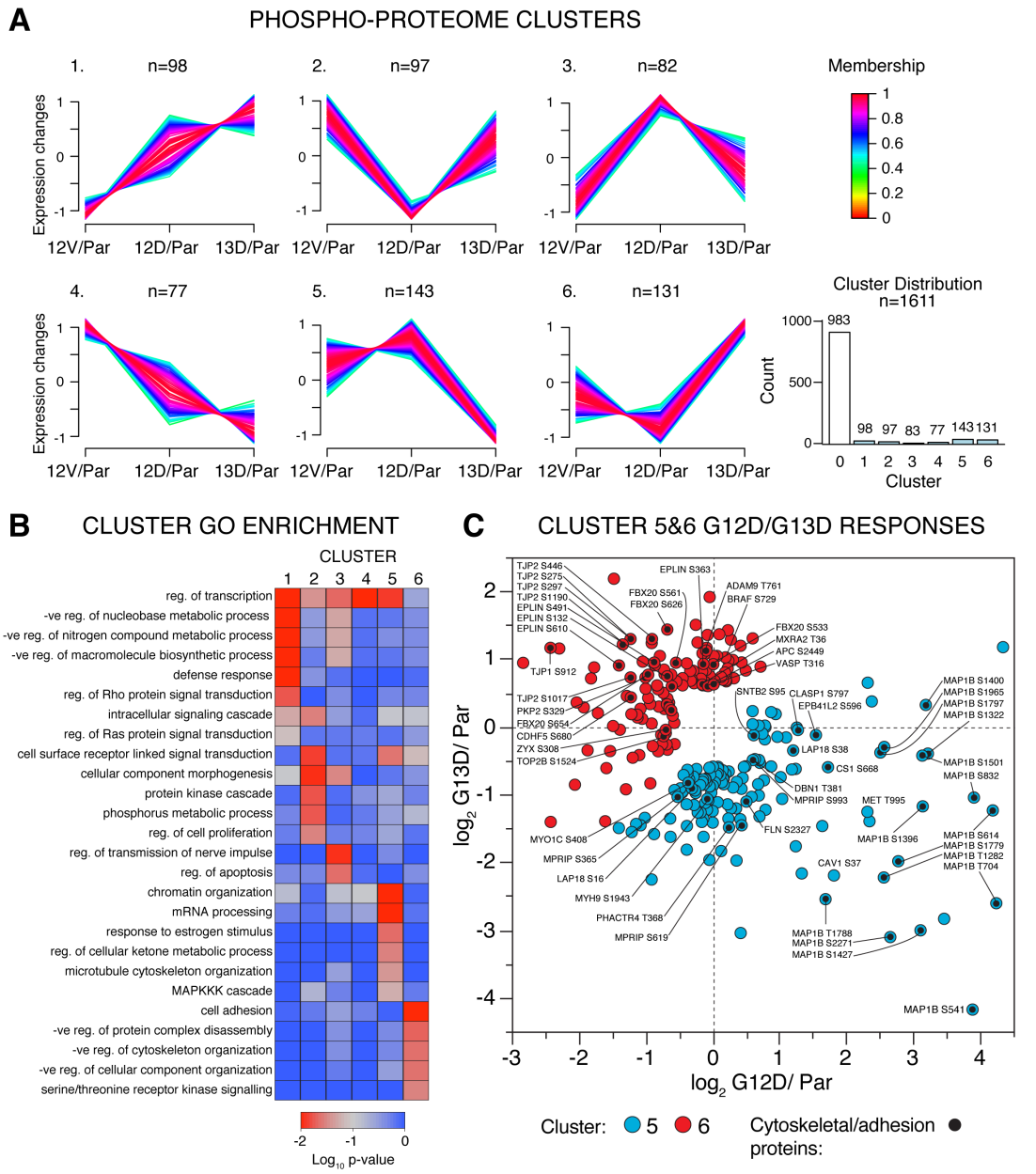


Fig. 3.30. Phosphopeptides displaying differential codon 12 and 13 mutant K-Ras responses. (**A**) Clustering of phosphopeptides based on response to G12V, G12D and G13D K-Ras mutations. Mutant/ Par ratios were subject to clustering with the Fuzzy c means algorithm using GProX. Clusters 5 and 6 contain proteins that are probable signatures of codon 12 and 13-specific signalling in colorectal cancer. (**B**) Gene Ontology (GO) analysis of each cluster performed using the DAVID functional annotation tool 6.7. (**C**) Scatterplot of clusters 5 and 6. Cytoskeletal and adhesion proteins are highlighted. GProX cluster analysis performed with the help of Dr. Dean Hammond (University of Liverpool).

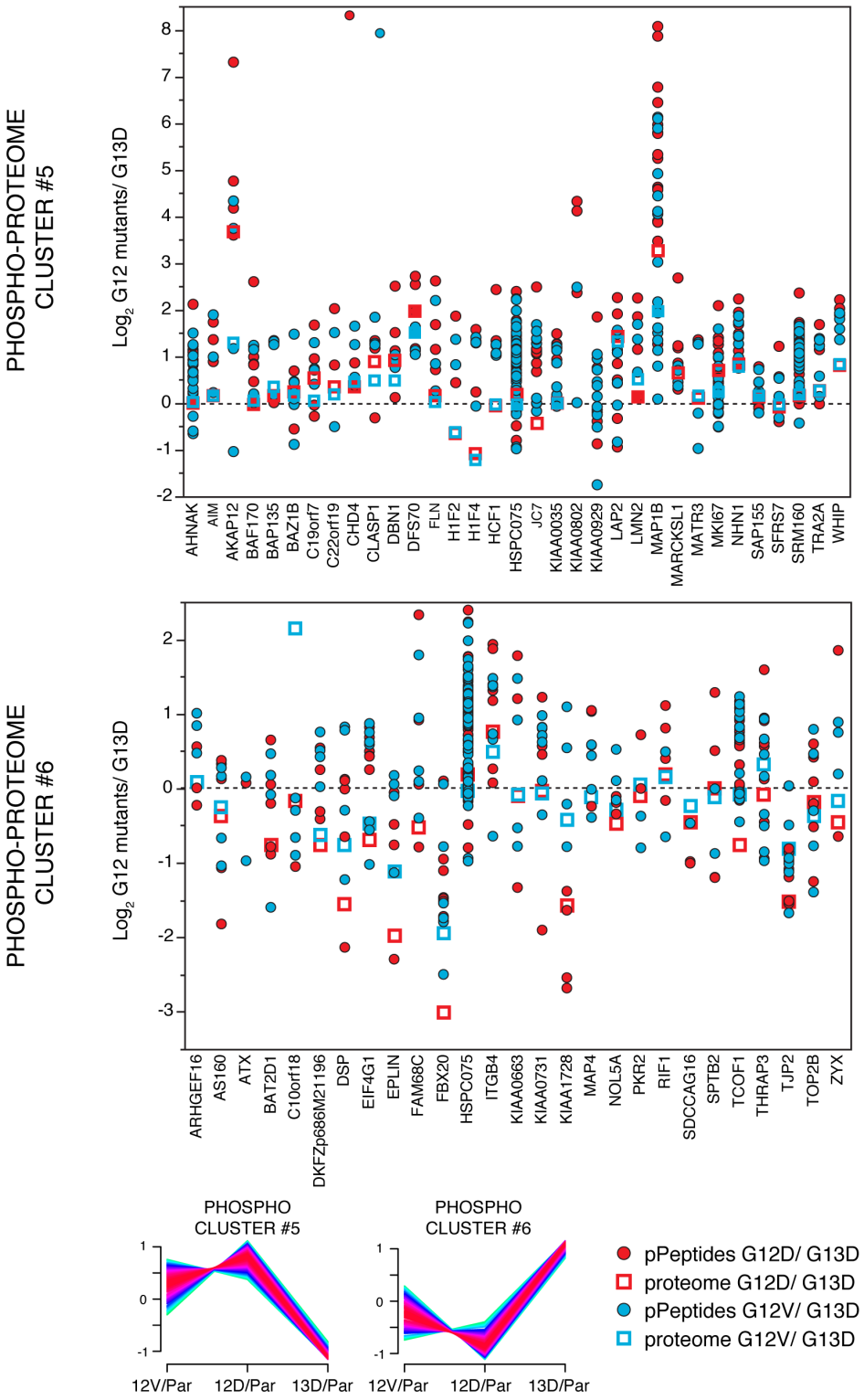


Fig. 3.31. Relationship between phosphopeptide response and proteome changes. Proteins represented by phosphopeptides from clusters 5 and 6 (see Fig. 3.30) that had ≥ 3 phosphopeptides identified in the entire dataset were curated, along with their respective protein abundance values.

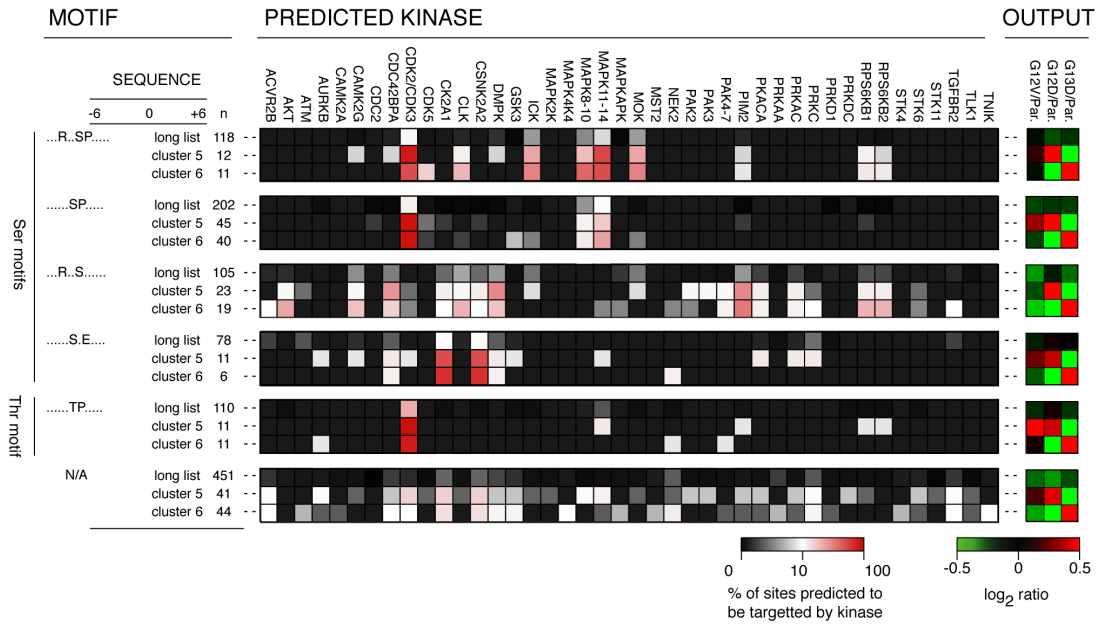


Fig. 3.32. Predicted kinase regulators of phosphosites associated with differential codon 12 and 13 mutant K-Ras responses. MotifX analysis identified over-represented linear phosphorylation motifs from those present in the entire phospho-proteome data set (long list) and phospho-proteome GProX clusters 5 and 6 (Chou and Schwartz, 2011; Schwartz and Gygi, 2005). NetworKIN was then utilised to identify candidate kinases (Linding et al., 2008). Output column indicates average responses of sites associated with each motif. n indicates the total number of over-represented motifs identified by MotifX and used in the output column. Figure generated with the aid of Prof. Ian Prior (University of Liverpool).

Given the central role of protein kinases in regulating Ras responses, their contribution to the dataset was examined. 38 kinases in total were identified, with 35 kinase phosphopeptides from 96 responsive (≥ 1.5-fold change compared to parental) to the presence of an oncogenic K-Ras mutation. Notable kinases identified include EGFR, MET, MAP2K2 (MEK2) and ERK2. The identified phosphosite of ERK2, T185, becomes phosphorylated during activation and was down-regulated in G13D versus codon 12 cell lines. To predict which kinases may be modulating the responses observed in phosphoproteome clusters 5 and 6, MotifX was utilised to identify over-represented phosphorylation motifs (Chou and Schwartz, 2011; Schwartz and Gygi, 2005), before NetworKIN analysis predicted the candidate kinases through integration of consensus substrate motifs and contextual modelling (Linding et al., 2008) (Fig. 3.32). Kinases involved in the regulation of the cell cycle and promotion of proliferation, e.g. cyclin-dependent, MAP and MOK kinases, were identified as potential kinases targeting members of clusters 5 and 6.

3.3.3 DCLK1 IS EXCLUSIVELY UPREGULATED IN SW48 ISOGENIC CELLS HARBOURING CODON 12 K-RAS MUTATIONS

Within the proteome dataset, doublecortin-like kinase-1 (DCLK1) and A-kinase anchor protein 12 (AKAP12) were prime examples of proteins responsive to K-RasG12D, showing ≥ 8-fold up-regulation versus Parental. Immunoblotting confirmed the distinctive expression pattern of these proteins (Fig. 3.33A), which was recapitulated in a second independent clone of each isogenic SW48 cell line (Fig. 3.34). Up-regulation of MET and Caveolin 1 was also confirmed in codon 12 mutant K-Ras cells (Fig. 3.33A). The observation of codon 12-specific DCLK1 up-regulation is notable, since recent data suggests that this protein is a colon cancer tumour stem cell marker (Nakanishi et al., 2013). Together, these observations led us to examine the extent of the codon 12-specific phenomenon, by analysing a wider panel of SW48 cell lines that harbour other K-Ras mutations not tested in the proteomic analysis. Interestingly, DCLK1 expression was significantly increased in all cells harbouring a codon 12 K-Ras mutation, while K-RasG13D and N-RasG12V had expression levels equivalent to Parental (Fig. 3.33B and C). The molecular weight markers and detected peptides/ phosphopeptides suggest that isoforms 3 and 4 of DCLK1 are the up-regulated proteins (Fig 3.33B and D), which contain an active kinase but lack the doublecortin-like domains enabling microtubule binding (Burgess and Reiner, 2001). MET exhibited the same expression pattern as DCLK1, possibly indicating a shared mechanism promoting up-regulation, while TJP2/ ZO-2 was shown to be preferentially up-regulated in the codon 12 mutant cell line only.

3.3.4 DISCUSSION

By applying large-scale quantitative proteomics to the SW48 isogenic cell line panel, an unprecedented insight into differential signals emanating from oncogenic G12D, G12V and G13D K-Ras mutants has been achieved, in a colorectal cancer setting. Our first and possibly most prominent observation is that K-Ras mutants have distinct signalling outputs. Since codon 12 mutants share greater overlap than codon 13 mutations on both proteome and phosphoproteome levels, it could be the positioning of the mutation has greater influence on the outputs of oncogenic K-Ras, rather than the type of amino acid substitution. This observation is important, as almost all Ras studies and Ras-related clinical trials assume that activating Ras mutations activate downstream signalling networks to the same extent.

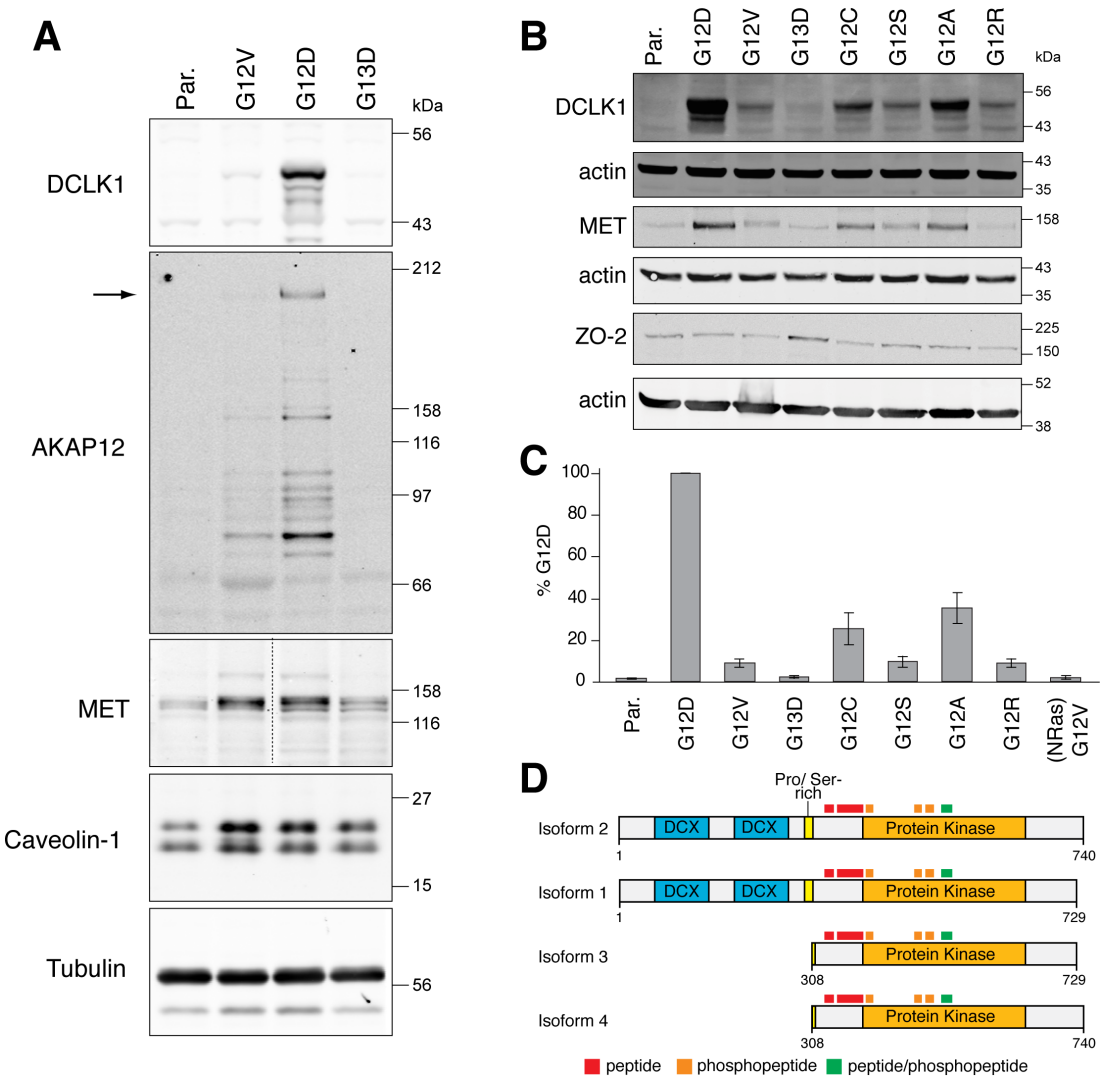


Fig. 3.33. DCLK1 is exclusively upregulated in cells harbouring codon 12 K-Ras mutations. (**A**) Immunoblot verification of selected hits from the MS-based SILAC analysis. (**B**) Immunoblot analysis of a wider panel of isogenic SW48 cell lines, harbouring K-Ras mutations not tested in the proteomic analysis, demonstrates that DCLK1 and MET are upregulated in the presence of a codon 12 K-Ras mutation, whereas ZO-2 is upregulated specifically by a codon 13 mutation. $n \geq 3$ for each panel. (**C**) Densitometry analysis DCLK1 expression across the entire SW48 cell line panel, including cells harbouring an N-RasG12V mutation. Bars represent mean ± SD. Signal normalised to G12D response. (**D**) Location of detected peptides and phosphopeptides along human DCLK1 isoforms. Part A includes immunoblots performed by Dr. Dean Hammond (University of Liverpool). ZO-2 and MET blots supplied by Dr. Emma Rusilowicz (University of Liverpool).

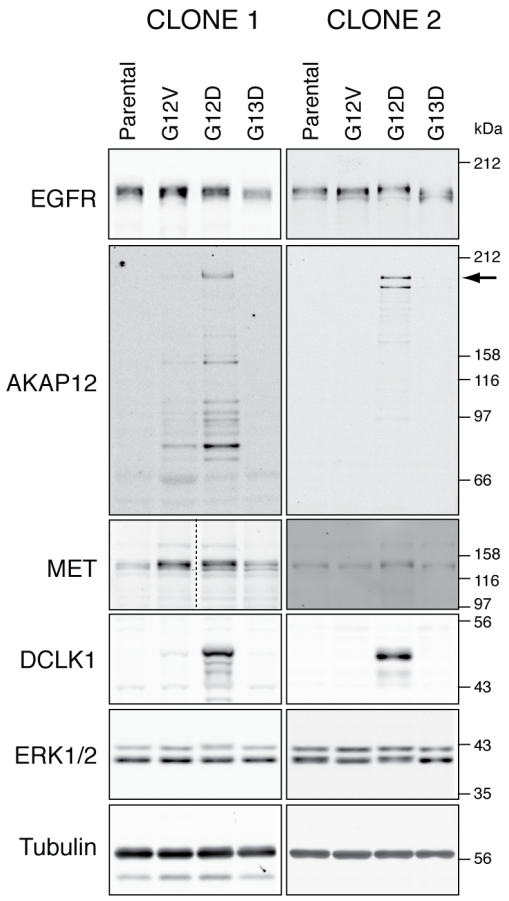


Fig. 3.34. Recapitulation of results in a second, independent SW48 clone. Both sets of independent clones demonstrate equivalent up-regulation of selected proteins. Immunoblots performed by Dr. Dean Hammond (University of Liverpool).

To understand the mechanisms that differentiate codon 12 and 13 mutant K-Ras signalling, 274 phosphopeptides and 115 proteins were identified that responded differently to the presence of the different codon mutations. Amongst these, AKAP12, a cell adhesion protein that orchestrates interactions between kinases (Akakura and Gelman, 2012; Malbon et al., 2004), and caveolin 1, an essential component of membrane caveolae that can modulate Ras signalling (Kranenburg et al., 2001; Parton and del Pozo, 2013), were specifically up-regulated in codon 12 mutant cells versus codon 13, with up-regulation of caveolin previously associated with oncogenic K-Ras (Roy et al., 2013). As these proteins are both potential tumour suppressors (Su et al., 2010), with AKAP12 commonly lost in CRC (Liu et al., 2011), it seems somewhat paradoxical that these proteins are up-regulated in the presence of oncogenic Ras. Yet, these responses could be the cell's attempt to dampen the distinct, aberrant signalling emanating from codon 12 mutant K-Ras. The HGF-receptor c-MET was similarly responsive to codon 12 mutations, but its up-regulation was also consistently observed in a wider panel cell lines harbouring

other codon 12 mutations. The up-regulation of MET has a well-established role in Ras-mediated carcinogenesis (Furge et al., 2001; Webb et al., 1998) and despite being upstream of Ras it could still be affected by oncogenic Ras signalling, since interplay between Ras pathway activation and growth factor receptors has been demonstrated by the inhibition of BRAF leading to the activation of EGFR via loss of negative feedback (Prahallad et al., 2012). In contrast, aldehyde dehydrogenase ALDH3A1 and TJP2 ZO-2 were selectively up-regulated in codon 13 mutant cell lines. ALDH3A1 is involved in the oxidation of a range of endogenous aldehydes and was originally designated as the tumour ALDH, due to its increased expression in a variety of cancers, and may contribute to drug resistance (Parajuli et al., 2014). TJP2 ZO-2 is a member of the membrane-associated guanylate kinase (MAGUK) family, which have a role as a scaffolding protein, as well as in the organisation of tight junctions and adherens junctions.

The most amplified protein in our dataset was DCLK1, which was most responsive to K-RasG12D but also consistently up-regulated in cells expressing codon 12 mutant K-Ras. DCLK2 has been previously identified as a synthetically lethal component in a genome-wide RNAi screen of colorectal DLD1 cells (Luo et al., 2009). Although the function of DCLK1 is poorly understood, consultation of the Cancer Cell Line Encyclopaedia, which contains data from over 1,000 cell lines (Barretina et al., 2012), revealed co-responsiveness with MAP1B, a major cytoskeletal protein and prominent phosphosite responder in codon 12 mutant cells (Lee et al., 2008), suggesting functional interaction between these oncogenic K-Ras responsive proteins. DCLK1 has an emerging role in colorectal carcinogenesis, as it has recently been identified as a CRC tumour stem cell-specific marker (Nakanishi et al., 2013), where ablation of DCLK1 positive tumour stem cells caused regression of polyps without damage to surrounding normal tissue. Plus, DCLK1 is frequently over-expressed in CRC and associated with a poorer prognosis/ overall survival (Gagliardi et al., 2012). Moreover, siRNA knockdown or inhibition of DCLK1 reduces colorectal and pancreatic proliferation (Weygant et al., 2014). As of yet, there has been no connection between DCLK1 expression and K-Ras mutational status, yet this association is likely to be highly context dependent since no significant correlation was identified in the Cancer Cell Line Encyclopaedia. Considering the size of DCLK1 isoform identified on immunoblots and the peptides/ phospho-peptides detected, the DCLK1 isoforms observed are the short isoforms 3 and 4.

These proteins lack the doublecortin domains necessary for microtubule binding and have a reported role in modulating memory and cognitive abilities (Francis et al., 1999; Le Hellard et al., 2009). Yet, it could be possible that DCLK1 isoform 1 or 2 is being converted into the shorter isoforms, through cleavage by calpain (Burgess and Reiner, 2001), but cleavage would have to be extremely rapid and incessant to prevent the observation of the longer DCLK1 isoforms. The pattern of expression of DCLK1, AKAP12 and MET were recapitulated in a second, independent clone, indicating that the observed results are not due to clonal selection, but represent responses attributable to oncogenic K-Ras signalling.

The data presented in this thesis provides the first unbiased, global screen of signals emanating from different oncogenic K-Ras mutations. The use of isogenic cell lines enables identification of specific signalling signatures without the confounding effects of differences in genetic background. One key observation of this study is that each Ras mutation induces a unique signalling response, meaning the precise Ras mutational status should be considered in experimental designs and interpretation of results when studying oncogenic Ras function. Several recent studies identified synthetically lethal genes when depleted/ inhibited in cell lines harbouring various K-Ras mutations (Barbie et al., 2009; Luo et al., 2009; Sarthy et al., 2007; Scholl et al., 2009; Singh et al., 2012; Steckel et al., 2012). Each study used different cell types that harboured a variety of K-Ras mutations and found inconsistent responses between cell lines. Not only are these inconsistent responses likely to arise due to differences in genetic backgrounds between cell lines, but the present data here also suggest they arise due to differential signalling from the different K-Ras mutations.

In summary, several key proteins with known functions related to colorectal carcinogenesis were shown to be differentially responsive to codon 12 or 13 K-Ras mutations. Amongst these is the colon cancer stem cell marker DCLK1. Most strikingly, we found that the three most abundant K-Ras mutations observed in colorectal cancer: G12D, G12V and G13D, generate distinct signalling signatures at both the proteome and phosphoproteome levels, showing that K-Ras mutations are not equal in their signalling outputs.

4 – OUTLOOK AND CONCLUSION

Up to now there has been uncertainty regarding the number of Ras molecules per cell and which isoform is the most abundant. Previous attempts to measure cellular Ras abundance utilise either an immunoprecipitation or SDS-PAGE step prior to addition of AQUA peptides (Halvey et al., 2012; Ruppen-Canas et al., 2012; Wang et al., 2011). Consequently, these methods are vulnerable to inaccuracies and variations, exemplified by the Wang and Ruppen-Canas studies obtaining ~100-fold different results despite employing the same Ras quantification technique. The SRM-based Ras quantification method and PSAQ standards developed in the present thesis has enabled the most accurate measurement of cellular Ras isoform abundance to date. The described Ras quantification technique is also easily implementable in most biological laboratories, since SDS-PAGE is a very familiar fractionation technique to these laboratories and the in-gel digest steps do not require specialist knowledge. The developed technique reliably detected H-, K(B)- and N-Ras at endogenous levels, as well as other peptides shared between isoforms to verify the isoform-specific quantification. Furthermore, the PSAQ technique accounts for variations in peptide recovery from gel pieces and incomplete digestion. Despite not being detected at endogenous levels using the 4000 QTRAP, the described K(A)-Ras-specific and elastase-derived total Ras peptides could be utilised in future studies employing newer and more sensitive instrumentation. As a result, the Ras quantification technique detailed in this thesis enables quantification of almost all aspects of cellular Ras abundance.

Over 250,000 Ras molecules were measured in a Parental SW48 cell, with K-Ras the most abundant isoform, meaning per μm^2 of plasma membrane ~114 Ras proteins are bound. Intriguingly, we measured a varying amount of Ras between isogenic SW48 cell lines harbouring a range of heterozygous Ras mutations. SW48 cells harbouring G12D and G13D K-Ras mutations, the first and third most common K-Ras mutations in CRC (Prior et al., 2012; Vaughn et al., 2011), contained almost double the amount of Ras compared to the Parental, with around 200,000 mutant K-Ras proteins expressed. In contrast, cells harbouring K-RasG12V, the second most common mutation in CRC, and those with H- and N-RasG12V, exhibited little difference for the Parental cell line, with ~30,000 copies of mutant protein. These differing amounts of Ras could be attributable to the differing transforming activities

of each Ras mutant. The G12V mutation could have such strong transforming activity that the cell is unable to tolerate many copies of Ras afflicted with this mutation, irrespective of which isoform is mutated. An insight into the possible additional potency of the G12V mutation came from immunoblot analyses of phospho-AKT, which showed, despite being expressed at lower levels than K-RasG12D and G13D, K-RasG12V was able to reduce the activation of the PI3K pathway to equivalent levels. Thus, the cell may be able to tolerate a certain amount of oncogenic Ras signalling up to a certain threshold before it limits expression of the mutant protein. The G12V Ras mutation may reach this threshold sooner than other mutations, possibly explaining why it is expressed at 6-fold lower levels compared to G12D and G13D K-Ras mutations. The range of abundances observed for other K-Ras mutations present in the isogenic panel could also reflect their respective transforming potentials. Since K-RasG12D is the most common Ras mutation in CRC, it could be that a large proportion of colon tumours prefer the least transforming Ras mutation during colorectal carcinogenesis, to limit the stress imposed on the cell from the aberrant signals emanating from the mutant Ras protein.

Large-scale proteomic and phospho-proteomic profiling of isogenic SW48 cell lines revealed that G12V, G12D and G13D K-Ras mutations revealed each activated a distinct signalling network. Yet, the codon 12 mutations (G12V and G12D K-Ras) exhibited the greatest degree of signalling overlap, suggesting that, although each mutation signals uniquely, the position of the mutation has a greater effect on signalling than an amino acid substitution. Several proteins relevant to colorectal carcinogenesis were found to be sensitive to certain Ras mutations. Of note is DCLK1, a marker of colon tumour stem cells and tumour-initiating cells in the pancreas (Bailey et al., 2014; Nakanishi et al., 2013), which was consistently up-regulated in SW48 cells harbouring codon 12 mutant K-Ras. This up-regulation seems unique to K-Ras, since N-RasG12V did not amplify DCLK1 expression. Although early Ras work revealed that different Ras mutations have a range of transforming activities (Der et al., 1986; Seeburg et al., 1984; Sloan et al., 1990), several recent studies do not take this into account (Barbie et al., 2009; Scholl et al., 2009; Steckel et al., 2012). The data presented in this thesis re-enforces this early Ras research, by showing that different Ras mutations induce a unique signalling response and selective up-regulation of different proteins.

5 – REFERENCES

Abankwa, D., Gorfe, A.A., and Hancock, J.F. (2007). Ras nanoclusters: molecular structure and assembly. Semin Cell Dev Biol *18*, 599-607.
ABSCIEX (2010). Hardware guide 4000 QTRAP System. *18*, 599-607.
Aguiar, M., Haas, W., Beausoleil, S.A., Rush, J., and Gygi, S.P. (2010). Gas-phase rearrangements do not affect site localization reliability in phosphoproteomics data sets. J Proteome Res *9*, 3103-3107.
Ahearn, I.M., Haigis, K., Bar-Sagi, D., and Philips, M.R. (2012). Regulating the regulator: post-translational modification of RAS. Nat Rev Mol Cell Biol *13*, 39-51.
Akakura, S., and Gelman, I.H. (2012). Pivotal Role of AKAP12 in the Regulation of Cellular Adhesion Dynamics: Control of Cytoskeletal Architecture, Cell Migration, and Mitogenic Signaling. J Signal Transduct *2012*, 529179.
Al-Mulla, F., Milner-White, E.J., Going, J.J., and Birnie, G.D. (1999). Structural differences between valine-12 and aspartate-12 Ras proteins may modify carcinoma aggression. J Pathol *187*, 433-438.
Aldridge, B.B., Burke, J.M., Lauffenburger, D.A., and Sorger, P.K. (2006). Physicochemical modelling of cell signalling pathways. Nat Cell Biol *8*, 1195-1203.
Alvarez-Moya, B., Barcelo, C., Tebar, F., Jaumot, M., and Agell, N. (2011). CaM interaction and Ser181 phosphorylation as new K-Ras signaling modulators. Small GTPases *2*, 99-103.
Anderson, L., and Hunter, C.L. (2006). Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. Mol Cell Proteomics *5*, 573-588.
Andersson, L., and Porath, J. (1986). Isolation of phosphoproteins by immobilized metal (Fe3+) affinity chromatography. Anal Biochem *154*, 250-254.
Andreyev, H.J., Norman, A.R., Cunningham, D., Oates, J., Dix, B.R., Iacopetta, B.J., Young, J., Walsh, T., Ward, R., Hawkins, N.*, et al.* (2001). Kirsten ras mutations in patients with colorectal cancer: the 'RASCAL II' study. Br J Cancer *85*, 692-696.
Andreyev, H.J., Norman, A.R., Cunningham, D., Oates, J.R., and Clarke, P.A. (1998). Kirsten ras mutations in patients with colorectal cancer: the multicenter "RASCAL" study. J Natl Cancer Inst *90*, 675-684.
Anwar, S., Frayling, I.M., Scott, N.A., and Carlson, G.L. (2004). Systematic review of genetic influences on the prognosis of colorectal cancer. Br J Surg *91*, 1275-1291.
Apolloni, A., Prior, I.A., Lindsay, M., Parton, R.G., and Hancock, J.F. (2000). H-ras but not K-ras traffics to the plasma membrane through the exocytic pathway. Mol Cell Biol *20*, 2475-2487.
Apweiler, R., Bairoch, A., Wu, C.H., Barker, W.C., Boeckmann, B., Ferro, S., Gasteiger, E., Huang, H., Lopez, R., Magrane, M.*, et al.* (2004). UniProt: the Universal Protein knowledgebase. Nucleic Acids Res *32*, D115-119.
Asara, J.M., Christofk, H.R., Freimark, L.M., and Cantley, L.C. (2008). A label-free quantification method by MS/MS TIC compared to SILAC and spectral counting in a proteomics screen. Proteomics *8*, 994-999.
Bailey, J.M., Alsina, J., Rasheed, Z.A., McAllister, F.M., Fu, Y.Y., Plentz, R., Zhang, H., Pasricha, P.J., Bardeesy, N., Matsui, W.*, et al.* (2014). DCLK1 marks a morphologically distinct subpopulation of cells with stem cell properties in preinvasive pancreatic cancer. Gastroenterology *146*, 245-256.
Baker, T.L., Zheng, H., Walker, J., Coloff, J.L., and Buss, J.E. (2003). Distinct rates of palmitate turnover on membrane-bound cellular and oncogenic H-ras. J Biol Chem *278*, 19292-19300.

Ballester, R., Furth, M.E., and Rosen, O.M. (1987). Thorbol ester- and protein kinase C-mediated phosphorylation of the cellular Kirsten ras gene product. J Bio Chem *262*, 2688-2695.
Bando, H., Yoshino, T., Yuki, S., Shinozaki, E., Nishina, T., Kadowaki, S., Yamazaki, K., Kajiura, S., Tsuchihara, K., Fujii, S.*, et al.* (2012). Clinical outcome of Japanese metastatic colorectal cancer patients harbouring the KRAS p.G13D mutation treated with cetuximab + irinotecan. Jpn J Clin Oncol *42*, 1146-1151.
Bantscheff, M., Lemeer, S., Savitski, M.M., and Kuster, B. (2012). Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present. Anal Bioanal Chem *404*, 939-965.
Bantscheff, M., Schirle, M., Sweetman, G., Rick, J., and Kuster, B. (2007). Quantitative mass spectrometry in proteomics: a critical review. Anal Bioanal Chem *389*, 1017-1031.
Barbie, D.A., Tamayo, P., Boehm, J.S., Kim, S.Y., Moody, S.E., Dunn, I.F., Schinzel, A.C., Sandy, P., Meylan, E., Scholl, C.*, et al.* (2009). Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. Nature *462*, 108-U122.
Barr, J.R., Maggio, V.L., Patterson, D.G., Jr., Cooper, G.R., Henderson, L.O., Turner, W.E., Smith, S.J., Hannon, W.H., Needham, L.L., and Sampson, E.J. (1996). Isotope dilution--mass spectrometric quantification of specific proteins: model application with apolipoprotein A-I. Clin Chem *42*, 1676-1682.
Barretina, J., Caponigro, G., Stransky, N., Venkatesan, K., Margolin, A.A., Kim, S., Wilson, C.J., Lehar, J., Kryukov, G.V., Sonkin, D.*, et al.* (2012). The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature *483*, 603-607.
Baselga, J. (2001). The EGFR as a target for anticancer therapy - focus on cetuximab. European Journal of Cancer *37*, S16-S22.
Baty, J.D., and Robinson, P.R. (1977). Single and multiple ion recording techniques for the analysis of diphenylhydantoin and its major metabolite in plasma. Biomed Mass Spectrom *4*, 36-41.
Bazan, V., Migliavacca, M., Zanna, I., Tubiolo, C., Grassi, N., Latteri, M.A., La Farina, M., Albanese, I., Dardanoni, G., Salerno, S.*, et al.* (2002). Specific codon 13 K-ras mutations are predictive of clinical outcome in colorectal cancer patients, whereas codon 12 K-ras mutations are associated with mucinous histotype. Ann Oncol *13*, 1438-1446.
Beausoleil, S.A., Jedrychowski, M., Schwartz, D., Elias, J.E., Villen, J., Li, J., Cohn, M.A., Cantley, L.C., and Gygi, S.P. (2004). Large-scale characterization of HeLa cell nuclear phosphoproteins. Proc Natl Acad Sci U S A *101*, 12130-12135.
Bentley, C., Jurinka, S.S., Kljavin, N.M., Vartanian, S., Ramani, S.R., Gonzalez, L.C., Yu, K., Modrusan, Z., Du, P., Bourgon, R.*, et al.* (2013). A requirement for wild-type Ras isoforms in mutant KRas-driven signalling and transformation. Biochem J *452*, 313-320.
Beynon, R.J., Doherty, M.K., Pratt, J.M., and Gaskell, S.J. (2005). Multiplexed absolute quantification in proteomics using artificial QCAT proteins of concatenated signature peptides. Nat Methods *2*, 587-589.
Biosciences, A. (1999). Reversed phase chromatography principles and methods. Amersham Biosciences Literature, 6-27.
Birktoft, J.J., and Breddam, K. (1994). Glutamyl endopeptidases. Methods Enzymol *244*, 114-126.
Bivona, T.G., Perez De Castro, I., Ahearn, I.M., Grana, T.M., Chiu, V.K., Lockyer, P.J., Cullen, P.J., Pellicer, A., Cox, A.D., and Philips, M.R. (2003). Phospholipase Cgamma activates Ras on the Golgi apparatus by means of RasGRP1. Nature *424*, 694-698.

Bivona, T.G., Quatela, S.E., Bodemann, B.O., Ahearn, I.M., Soskis, M.J., Mor, A., Miura, J., Wiener, H.H., Wright, L., Saba, S.G.*, et al.* (2006). PKC regulates a farnesyl-electrostatic switch on K-Ras that promotes its association with Bcl-XL on mitochondria and induces apoptosis. Mol Cell *21*, 481-493.
Boersema, P.J., Mohammed, S., and Heck, A.J. (2009). Phosphopeptide fragmentation and analysis by mass spectrometry. J Mass Spectrom *44*, 861-878.
Boisvert, F.M., Ahmad, Y., Gierlinski, M., Charriere, F., Lamont, D., Scott, M., Barton, G., and Lamond, A.I. (2012). A quantitative spatial proteomics analysis of proteome turnover in human cells. Mol Cell Proteomics *11*, M111 011429.
Bollag, G., and McCormick, F. (1991). Differential regulation of rasGAP and neurofibromatosis gene product activities. Nature *351*, 576-579.
Bondarenko, P.V., Chelius, D., and Shaler, T.A. (2002). Identification and relative quantitation of protein mixtures by enzymatic digestion followed by capillary reversed-phase liquid chromatography-tandem mass spectrometry. Anal Chem *74*, 4741-4749.
Bordbar, A., Monk, J.M., King, Z.A., and Palsson, B.O. (2014). Constraint-based models predict metabolic and associated cellular functions. Nat Rev Genet *15*, 107-120.
Boriack-Sjodin, P.A., Margarit, S.M., Bar-Sagi, D., and Kuriyan, J. (1998). The structural basis of the activation of Ras by Sos. Nature *394*, 337-343.
Bos, J.L. (1989). ras oncogenes in human cancer: a review. Cancer Res *49*, 4682-4689.
Bos, J.L., Fearon, E.R., Hamilton, S.R., Verlaan-de Vries, M., van Boom, J.H., van der Eb, A.J., and Vogelstein, B. (1987). Prevalence of ras gene mutations in human colorectal cancers. Nature *327*, 293-297.
Bos, J.L., Rehmann, H., and Wittinghofer, A. (2007). GEFs and GAPs: critical elements in the control of small G proteins. Cell *129*, 865-877.
Bos, J.L., Toksoz, D., Marshall, C.J., Verlaan-de Vries, M., Veeneman, G.H., van der Eb, A.J., van Boom, J.H., Janssen, J.W.G., and Steenvoorden, A.C.M. (1985). Amino-acid substitutions at codon 13 of the N-ras oncogene in human acute myeloid leukaemia. Nature *315*, 726-730.
Boyartchuk, V.L., Ashby, M.N., and Rine, J. (1997). Modulation of Ras and a-factor function by carboxyl-terminal proteolysis. Science *275*, 1796-1800.
Bremner, R., and Balmain, A. (1990). Genetic Changes in Skin Tumor Progression - Correlation between Presence of a Mutant Ras Gene and Loss of Heterozygosity on Mouse Chromosome-7. Cell *61*, 407-417.
Brennan, D.F., Dar, A.C., Hertz, N.T., Chao, W.C., Burlingame, A.L., Shokat, K.M., and Barford, D. (2011). A Raf-induced allosteric transition of KSR stimulates phosphorylation of MEK. Nature *472*, 366-369.
Brun, V., Masselon, C., Garin, J., and Dupuis, A. (2009). Isotope dilution strategies for absolute quantitative proteomics. J Proteomics *72*, 740-749.
Buhrman, G., Holzapfel, G., Fetics, S., and Mattos, C. (2010). Allosteric modulation of Ras positions Q61 for a direct role in catalysis. Proc Natl Acad Sci U S A *107*, 4931-4936.
Buhrman, G., Kumar, V.S., Cirit, M., Haugh, J.M., and Mattos, C. (2011). Allosteric modulation of Ras-GTP is linked to signal transduction through RAF kinase. J Biol Chem *286*, 3323-3331.
Buhrman, G., Wink, G., and Mattos, C. (2007). Transformation efficiency of RasQ61 mutants linked to structural features of the switch regions in the presence of Raf. Structure *15*, 1618-1629.
Burgess, H.A., and Reiner, O. (2001). Cleavage of doublecortin-like kinase by calpain releases an active kinase fragment from a microtubule anchorage domain. J Biol Chem *276*, 36397-36403.

Burlet, O., Yang, C.Y., and Gaskell, S.J. (1992). Influence of cysteine to cysteic acid oxidation on the collision-activated decomposition of protonated peptides: Evidence for intraionic interactions. J Am Soc Mass Spectrom *3*, 337-344.
Capella, G., Cronauer-Mitra, S., Pienado, M.A., and Perucho, M. (1991). Frequency and spectrum of mutations at codons 12 and 13 of the c-K-ras gene in human tumors. Environ Health Perspect *93*, 125-131.
Capon, D.J., Chen, E.Y., Levinson, A.D., Seeburg, P.H., and Goeddel, D.V. (1983a). Complete nucleotide sequences of the T24 human bladder carcinoma oncogene and its normal homologue. Nature *302*, 33-37.
Capon, D.J., Seeburg, P.H., McGrath, J.P., Hayflick, J.S., Edman, U., Levinson, A.D., and Goeddel, D.V. (1983b). Activation of Ki-ras2 gene in human colon and lung carcinomas by two different point mutations. Nature *304*, 507-513.
Casey, P.J., Solski, P.A., Der, C.J., and Buss, J.E. (1989). p21ras is modified by a farnesyl isoprenoid. Proc Natl Acad Sci U S A *86*, 8323-8327.
Castellano, E., and Downward, J. (2010). Role of RAS in the regulation of PI 3-kinase. Curr Top Microbiol Immunol *346*, 143-169.
Castellano, E., and Santos, E. (2011). Functional specificity of ras isoforms: so similar but so different. Genes Cancer *2*, 216-231.
Cepero, V., Sierra, J.R., Corso, S., Ghiso, E., Casorzo, L., Perera, T., Comoglio, P.M., and Giordano, S. (2010). MET and KRAS gene amplification mediates acquired resistance to MET tyrosine kinase inhibitors. Cancer Res *70*, 7580-7590.
Cerottini, J.P., Caplin, S., Saraga, E., Givel, J.C., and Benhattar, J. (1998). The type of K-ras mutation determines prognosis in colorectal cancer. Am J Surg *175*, 198-202.
Cespedes, M.V., Sancho, F.J., Guerrero, S., Parreno, M., Casanova, I., Pavon, M.A., Marcuello, E., Trias, M., Cascante, M., Capella, G.*, et al.* (2006). K-ras Asp12 mutant neither interacts with Raf, nor signals through Erk and is less tumorigenic than K-ras Val12. Carcinogenesis *27*, 2190-2200.
Chalkley, R.J., and Clauser, K.R. (2012). Modification site localization scoring: strategies and performance. Mol Cell Proteomics *11*, 3-14.
Chandra, A., Grecco, H.E., Pisupati, V., Perera, D., Cassidy, L., Skoulidis, F., Ismail, S.A., Hedberg, C., Hanzal-Bayer, M., Venkitaraman, A.R.*, et al.* (2012). The GDI-like solubilizing factor PDEdelta sustains the spatial organization and signalling of Ras family proteins. Nat Cell Biol *14*, 148-158.
Chang, E.H., Furth, M.E., Scolnick, E.M., and Lowy, D.R. (1982a). Tumorigenic transformation of mammalian cells induced by a normal human gene homologous to the oncogene of Harvey murine sarcoma virus. Nature *297*, 479-483.
Chang, E.H., Gonda, M.A., Ellis, R.W., Scolnick, E.M., and Lowy, D.R. (1982b). Human genome contains four genes homologous to transforming genes of Harvey and Kirsten murine sarcoma viruses. Proc Natl Acad Sci U S A *79*, 4848-4852.
Chang, Y.C., Lin, S.Y., Liang, S.Y., Pan, K.T., Chou, C.C., Chen, C.H., Liao, C.L., Khoo, K.H., and Meng, T.C. (2008). Tyrosine phosphoproteomics and identification of substrates of protein tyrosine phosphatase dPTP61F in Drosophila S2 cells by mass spectrometry-based substrate trapping strategy. J Proteome Res *7*, 1055-1066.
Chen, Z., Otto, J.C., Bergo, M.O., Young, S.G., and Casey, P.J. (2000). The C-terminal polylysine region and methylation of K-Ras are critical for the interaction between K-Ras and microtubules. J Biol Chem *275*, 41251-41257.
Cherfils, J., and Zeghouf, M. (2013). Regulation of small GTPases by GEFs, GAPs, and GDIs. Physiol Rev *93*, 269-309.
Chesa, P.G., Rettig, W.J., Melamed, M.R., Old, L.J., and Niman, H.L. (1987). Expression of p21ras in normal and malignant human tissues: lack of association with proliferation and malignancy. Proc Natl Acad Sci U S A *84*, 3234-3238.

Chicooree, N., Unwin, R.D., and Griffiths, J.R. (2014). The application of targeted mass spectrometry-based strategies to the detection and localization of post-translational modifications. Mass Spectrom Rev, DOI: 10.1002-mas.21421.
Chiu, V.K., Bivona, T., Hach, A., Sajous, J.B., Silletti, J., Wiener, H., Johnson, R.L., 2nd, Cox, A.D., and Philips, M.R. (2002). Ras signalling on the endoplasmic reticulum and the Golgi. Nat Cell Biol *4*, 343-350.
Chou, M.F., and Schwartz, D. (2011). Biological sequence motif discovery using motif-x. Curr Protoc Bioinformatics *Chapter 13*, Unit 13 15-24.
Choy, E., Chiu, V.K., Silletti, J., Feoktistov, M., Morimoto, T., Michaelson, D., Ivanov, I.E., and Philips, M.R. (1999). Endomembrane trafficking of ras: the CAAX motif targets proteins to the ER and Golgi. Cell *98*, 69-80.
Christensen, G.L., Kelstrup, C.D., Lyngso, C., Sarwar, U., Bogebo, R., Sheikh, S.P., Gammeltoft, S., Olsen, J.V., and Hansen, J.L. (2010). Quantitative phosphoproteomics dissection of seven-transmembrane receptor signaling using full and biased agonists. Mol Cell Proteomics *9*, 1540-1553.
Claperon, A., and Therrien, M. (2007). KSR and CNK: two scaffolds regulating RAS-mediated RAF activation. Oncogene *26*, 3143-3158.
Clyde-Smith, J., Silins, G., Gartside, M., Grimmond, S., Etheridge, M., Apolloni, A., Hayward, N., and Hancock, J.F. (2000). Characterization of RasGRP2, a plasma membrane-targeted, dual specificity Ras/Rap exchange factor. J Biol Chem *275*, 32260-32267.
Cohen, J.B., Broz, S.D., and Levinson, A.D. (1989). Expression of the H-ras proto-oncogene is controlled by alternative splicing. Cell *58*, 461-472.
Cohen, J.B., and Levinson, A.D. (1988). A point mutation in the last intron responsible for increased expression and transforming activity of the c-Ha-ras oncogene. Nature *334*, 119-124.
Cohen, P. (2002). The origins of protein phosphorylation. Nat Cell Biol *4*, E127-130.
Colby, W.W., Hayflick, J.S., Clark, S.G., and Levinson, A.D. (1986). Biochemical characterization of polypeptides encoded by mutated human Ha-ras1 genes. Mol Cell Biol *6*, 730-734.
Comisarow, M.B., and Marshall, A.G. (1974). Fourier transform ion cyclotron resonance spectroscopy. Chemical Physics Letters *25*, 282-283.
Cordero, M.M., Houser, J.J., and Wesdemiotis, C. (1993). The neutral products formed during backbone fragmentations of protonated peptides in tandem mass spectrometry. Anal Chem *65*, 1594-1601.
Cox, A.D., and Der, C.J. (2010). Ras history: The saga continues. Small GTPases *1*, 2-27.
Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol *26*, 1367-1372.
Cox, J., Neuhauser, N., Michalski, A., Scheltema, R.A., Olsen, J.V., and Mann, M. (2011). Andromeda: a peptide search engine integrated into the MaxQuant environment. J Proteome Res *10*, 1794-1805.
Dai, Q., Choy, E., Chiu, V., Romano, J., Slivka, S.R., Steitz, S.A., Michaelis, S., and Philips, M.R. (1998). Mammalian prenylcysteine carboxyl methyltransferase is in the endoplasmic reticulum. J Biol Chem *273*, 15030-15034.
de Hoffmann, E., and Stroobant, V. (2013). Mass Spectrometry (John Wiley & Sons).
De Roock, W., Jonker, D.J., Di Nicolantonio, F., Sartore-Bianchi, A., Tu, D., Siena, S., Lamba, S., Arena, S., Frattini, M., Piessevaux, H.*, et al.* (2010). Association of KRAS p.G13D mutation with outcome in patients with chemotherapy-refractory metastatic colorectal cancer treated with cetuximab. JAMA *304*, 1812-1820.

De Roock, W., Piessevaux, H., De Schutter, J., Janssens, M., De Hertogh, G., Personeni, N., Biesmans, B., Van Laethem, J.L., Peeters, M., Humblet, Y.*, et al.* (2008). KRAS wild-type state predicts survival and is associated to early radiological response in metastatic colorectal cancer treated with cetuximab. Annals of Oncology *19*, 508-515.
de Vos, A., Tong, L., Milburn, M., Matias, P., Jancarik, J., Noguchi, S., Nishimura, S., Miura, K., Ohtsuka, E., and Kim, S. (1988). Three-dimensional structure of an oncogene protein: catalytic domain of human c-H-ras p21. Science *239*, 888-893.
DeFeo, D., Gonda, M.A., Young, H.A., Chang, E.H., Lowy, D.R., Scolnick, E.M., and Ellis, R.W. (1981). Analysis of two divergent rat genomic clones homologous to the transforming gene of Harvey murine sarcoma virus. Proc Natl Acad Sci U S A *78*, 3328-3332.
DeGnore, J.P., and Qin, J. (1998). Fragmentation of phosphopeptides in an ion trap mass spectrometer. J Am Soc Mass Spectrom *9*, 1175-1188.
Der, C.J., Finkel, T., and Cooper, G.M. (1986). Biological and biochemical properties of human rasH genes mutated at codon 61. Cell *44*, 167-176.
Der, C.J., Krontiris, T.G., and Cooper, G.M. (1982). Transforming genes of human bladder and lung carcinoma cell lines are homologous to the ras genes of Harvey and Kirsten sarcoma viruses. Proc Natl Acad Sci U S A *79*, 3637-3640.
Desiere, F., Deutsch, E.W., King, N.L., Nesvizhskii, A.I., Mallick, P., Eng, J., Chen, S., Eddes, J., Loevenich, S.N., and Aebersold, R. (2006). The PeptideAtlas project. Nucleic Acids Res *34*, D655-658.
DeSouza, L.V., Taylor, A.M., Li, W., Minkoff, M.S., Romaschin, A.D., Colgan, T.J., and Siu, K.W. (2008). Multiple reaction monitoring of mTRAQ-labeled peptides enables absolute quantification of endogenous levels of a potential cancer marker in cancerous and normal endometrial tissues. J Proteome Res *7*, 3525-3534.
Dhillon, A.S., Hagan, S., Rath, O., and Kolch, W. (2007). MAP kinase signalling pathways in cancer. Oncogene *26*, 3279-3290.
Di Nicolantonio, F., Arena, S., Gallicchio, M., Zecchin, D., Martini, M., Flonta, S.E., Stella, G.M., Lamba, S., Cancelliere, C., Russo, M.*, et al.* (2008). Replacement of normal with mutant alleles in the genome of normal human cells unveils mutation-specific drug responses. Proc Natl Acad Sci U S A *105*, 20864-20869.
Diaz, L.A., Jr., Williams, R.T., Wu, J., Kinde, I., Hecht, J.R., Berlin, J., Allen, B., Bozic, I., Reiter, J.G., Nowak, M.A.*, et al.* (2012). The molecular evolution of acquired resistance to targeted EGFR blockade in colorectal cancers. Nature *486*, 537-540.
Doerr, A. (2008). Top-down mass spectrometry. Nature Methods *5*, 24-24.
Dole, M., Mack, L.L., Hines, R.L., Mobley, R.C., Ferguson, L.D., and Alice, M.B. (1968). Molecular Beams of Macroions. Journal of Chemical Physics *49*, 2240-2240.
Dongre, A.R., Jones, J.L., Somogyi, A., and Wysocki, V.H. (1996). Influence of peptide composition, gas-phase basicity, and chemical modification on fragmentation efficiency: Evidence for the mobile proton model. Journal of the American Chemical Society *118*, 8365-8374.
Douglas, D.J. (2009). Linear quadrupoles in mass spectrometry. Mass Spectrom Rev *28*, 937-960.
Douglas, D.J., Frank, A.J., and Mao, D. (2005). Linear ion traps in mass spectrometry. Mass Spectrom Rev *24*, 1-29.
Drapeau, G.R., Boily, Y., and Houmard, J. (1972). Purification and properties of an extracellular protease of Staphylococcus aureus. J Biol Chem *247*, 6720-6726.
Dulla, K., Daub, H., Hornberger, R., Nigg, E.A., and Korner, R. (2010). Quantitative site-specific phosphorylation dynamics of human protein kinases during mitotic progression. Mol Cell Proteomics *9*, 1167-1181.

Dupuis, A., Hennekinne, J.A., Garin, J., and Brun, V. (2008). Protein Standard Absolute Quantification (PSAQ) for improved investigation of staphylococcal food poisoning outbreaks. Proteomics *8*, 4633-4636.
Echols, N., Harrison, P., Balasubramanian, S., Luscombe, N.M., Bertone, P., Zhang, Z., and Gerstein, M. (2002). Comprehensive analysis of amino acid and nucleotide composition in eukaryotic genomes, comparing genes and pseudogenes. Nucleic Acids Res *30*, 2515-2523.
Eckhart, W., Hutchinson, M.A., and Hunter, T. (1979). An activity phosphorylating tyrosine in polyoma T antigen immunoprecipitates. Cell *18*, 925-933.
Eisfeld, A.K., Schwind, S., Hoag, K.W., Walker, C.J., Liyanarachchi, S., Patel, R., Huang, X., Markowitz, J., Duan, W., Otterson, G.A.*, et al.* (2014). NRAS isoforms differentially affect downstream pathways, cell growth, and cell transformation. Proc Natl Acad Sci U S A *111*, 4179-4184.
Ellermann, V., and Bang, O. (1909). Experimentelle Leukämie bei Hühnern. II. Zeitschrift für Hygiene und Infektionskrankheiten *63*, 231-272.
Ellis, R.W., Defeo, D., Shih, T.Y., Gonda, M.A., Young, H.A., Tsuchida, N., Lowy, D.R., and Scolnick, E.M. (1981). The p21 src genes of Harvey and Kirsten sarcoma viruses originate from divergent members of a family of normal vertebrate genes. Nature *292*, 506-511.
Emerson, S.D., Madison, V.S., Palermo, R.E., Waugh, D.S., Scheffler, J.E., Tsao, K.L., Kiefer, S.E., Liu, S.P., and Fry, D.C. (1995). Solution structure of the Ras-binding domain of c-Raf-1 and identification of its Ras interaction surface. Biochemistry *34*, 6911-6918.
Eng, J.K., McCormack, A.L., and Yates, J.R. (1994). An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. J Am Soc Mass Spectrom *5*, 976-989.
Escher, C., Reiter, L., MacLean, B., Ossola, R., Herzog, F., Chilton, J., MacCoss, M.J., and Rinner, O. (2012). Using iRT, a normalized retention time for more targeted measurement of peptides. Proteomics *12*, 1111-1121.
Esteban, L.M., Vicario-Abejon, C., Fernandez-Salguero, P., Fernandez-Medarde, A., Swaminathan, N., Yienger, K., Lopez, E., Malumbres, M., McKay, R., Ward, J.M.*, et al.* (2001). Targeted genomic disruption of H-ras and N-ras, individually or in combination, reveals the dispensability of both loci for mouse growth and development. Mol Cell Biol *21*, 1444-1452.
Fasano, O., Aldrich, T., Tamanoi, F., Taparowsky, E., Furth, M., and Wigler, M. (1984). Analysis of the transforming potential of the human H-ras gene by random mutagenesis. Proc Natl Acad Sci U S A *81*, 4008-4012.
Fearon, E.R. (2011). Molecular genetics of colorectal cancer. Annu Rev Pathol *6*, 479-507.
Fearon, E.R., and Vogelstein, B. (1990). A genetic model for colorectal tumorigenesis. Cell *61*, 759-767.
Fehrenbacher, N., Bar-Sagi, D., and Philips, M. (2009). Ras/MAPK signaling from endomembranes. Mol Oncol *3*, 297-307.
Fenn, J.B. (1993). Ion formation from charged droplets: Roles of geometry, energy, and time. J Am Soc Mass Spectrom *4*, 524-535.
Ferro, E., and Trabalzini, L. (2010). RalGDS family members couple Ras to Ral signalling and that's not all. Cell Signal *22*, 1804-1810.
Ficarro, S.B., McCleland, M.L., Stukenberg, P.T., Burke, D.J., Ross, M.M., Shabanowitz, J., Hunt, D.F., and White, F.M. (2002). Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae. Nat Biotechnol *20*, 301-305.

Fiorucci, G., and Hall, A. (1988). All three human ras genes are expressed in a wide range of tissues. Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression *950*, 81-83.
Fivaz, M., and Meyer, T. (2005). Reversible intracellular translocation of KRas but not HRas in hippocampal neurons regulated by Ca2+/calmodulin. J Cell Biol *170*, 429-441.
Francis, F., Koulakoff, A., Boucher, D., Chafey, P., Schaar, B., Vinet, M., Friocourt, G., McDonnell, N., Reiner, O., and Kahn, A. (1999). Doublecortin Is a Developmentally Regulated, Microtubule-Associated Protein Expressed in Migrating and Differentiating Neurons. Neuron *23*, 247-256.
Franken, S.M., Scheidig, A.J., Krengel, U., Rensland, H., Lautwein, A., Geyer, M., Scheffzek, K., Goody, R.S., Kalbitzer, H.R., Pai, E.F.*, et al.* (1993). Three-dimensional structures and properties of a transforming and a nontransforming glycine-12 mutant of p21H-ras. Biochemistry *32*, 8411-8420.
Frattini, M., Saletti, P., Romagnani, E., Martin, V., Molinari, F., Ghisletta, M., Camponovo, A., Etienne, L.L., Cavalli, F., and Mazzucchelli, L. (2007). PTEN loss of expression predicts cetuximab efficacy in metastatic colorectal cancer patients. Br J Cancer *97*, 1139-1145.
Furge, K.A., Kiewlich, D., Le, P., Vo, M.N., Faure, M., Howlett, A.R., Lipson, K.E., Woude, G.F.V., and Webb, C.P. (2001). Suppression of Ras-mediated tumorigenicity and metastasis through inhibition of the Met receptor tyrosine kinase. Proceedings of the National Academy of Sciences of the United States of America *98*, 10722-10727.
Furth, M.E., Aldrich, T.H., and Cordon-Cardo, C. (1987). Expression of ras proto-oncogene proteins in normal human tissues. Oncogene *1*, 47-58.
Futschik, M.E., and Carlisle, B. (2005). Noise-robust soft clustering of gene expression time-course data. J Bioinform Comput Biol *3*, 965-988.
Gagliardi, G., Goswami, M., Passera, R., and Bellows, C.F. (2012). DCLK1 immunoreactivity in colorectal neoplasia. Clin Exp Gastroenterol *5*, 35-42.
Gajate, P., Sastre, J., Bando, I., Alonso, T., Cillero, L., Sanz, J., Caldes, T., and Diaz-Rubio, E. (2012). Influence of KRAS p.G13D mutation in patients with metastatic colorectal cancer treated with cetuximab. Clin Colorectal Cancer *11*, 291-296.
Gallien, S., Duriez, E., and Domon, B. (2011). Selected reaction monitoring applied to proteomics. J Mass Spectrom *46*, 298-312.
Geiger, T., and Clarke, S. (1987). Deamidation, Isomerization, and Racemization at Asparaginyl and Aspartyl Residues in Peptides - Succinimide-Linked Reactions That Contribute to Protein-Degradation. Journal of Biological Chemistry *262*, 785-794.
Gerber, S.A., Rush, J., Stemman, O., Kirschner, M.W., and Gygi, S.P. (2003). Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. Proc Natl Acad Sci U S A *100*, 6940-6945.
Geyer, M., Herrmann, C., Wohlgemuth, S., Wittinghofer, A., and Kalbitzer, H.R. (1997). Structure of the Ras-binding domain of RalGEF and implications for Ras binding and signalling. Nat Struct Biol *4*, 694-699.
Geyer, M., Schweins, T., Herrmann, C., Prisner, T., Wittinghofer, A., and Kalbitzer, H.R. (1996). Conformational transitions in p21ras and in its complexes with the effector protein Raf-RBD and the GTPase activating protein GAP. Biochemistry *35*, 10308-10320.
Ghesquiere, B., and Gevaert, K. (2014). Proteomics methods to study methionine oxidation. Mass Spectrom Rev *33*, 147-156.
Gibbs, J.B., Sigal, I.S., Poe, M., and Scolnick, E.M. (1984). Intrinsic GTPase activity distinguishes normal and oncogenic ras p21 molecules. Proc Natl Acad Sci U S A *81*, 5704-5708.

Gideon, P., John, J., French, M., Lautwein, A., Clark, R., Scheffler, J.E., and Wittinghofer, A. (1992). Mutational and kinetic analyses of the GTPase-activating protein (GAP)-p21 interaction: the C-terminal domain of GAP is not sufficient for full activity. Mol Cell Biol *12*, 2050-2056.
Gill, S.C., and von Hippel, P.H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. Anal Biochem *182*, 319-326.
Gillet, L.C., Navarro, P., Tate, S., Rost, H., Selevsek, N., Reiter, L., Bonner, R., and Aebersold, R. (2012). Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. Mol Cell Proteomics *11*, O111 016717.
Gish, W., and States, D.J. (1993). Identification of protein coding regions by database similarity search. Nat Genet *3*, 266-272.
Goldfarb, M., Shimizu, K., Perucho, M., and Wigler, M. (1982). Isolation and preliminary characterization of a human transforming gene from T24 bladder carcinoma cells. Nature *296*, 404-409.
Goldfinger, L.E., Ptak, C., Jeffery, E.D., Shabanowitz, J., Han, J., Haling, J.R., Sherman, N.E., Fox, J.W., Hunt, D.F., and Ginsberg, M.H. (2007). An experimentally derived database of candidate Ras-interacting proteins. J Proteome Res *6*, 1806-1811.
Gong, C.X., Liu, F., Grundke-Iqbal, I., and Iqbal, K. (2006). Dysregulation of protein phosphorylation/dephosphorylation in Alzheimer's disease: a therapeutic target. J Biomed Biotechnol *2006*, 31825.
Goodwin, J.S., Drake, K.R., Rogers, C., Wright, L., Lippincott-Schwartz, J., Philips, M.R., and Kenworthy, A.K. (2005). Depalmitoylated Ras traffics to and from the Golgi complex via a nonvesicular pathway. J Cell Biol *170*, 261-272.
Grabocka, E., Pylayeva-Gupta, Y., Jones, M.J., Lubkov, V., Yemanaberhan, E., Taylor, L., Jeng, H.H., and Bar-Sagi, D. (2014). Wild-type H- and N-Ras promote mutant K-Ras-driven tumorigenesis by modulating the DNA damage response. Cancer Cell *25*, 243-256.
Graham, J., Muhsin, M., and Kirkpatrick, P. (2004). Cetuximab. Nat Rev Drug Discov *3*, 549-550.
Greggio, E. (2012). Role of LRRK2 kinase activity in the pathogenesis of Parkinson's disease. Biochem Soc Trans *40*, 1058-1062.
Guerrero, S., Casanova, I., Farré, L., Mazo, A., Capellà, G., and Mangues, R. (2000). K-ras Codon 12 Mutation Induces Higher Level of Resistance to Apoptosis and Predisposition to Anchorage-independent Growth Than Codon 13 Mutation or Proto-Oncogene Overexpression. Cancer Res *60*, 6750-6756.
Guerrero, S., Figueras, A., Casanova, I., Farre, L., Lloveras, B., Capella, G., Trias, M., and Mangues, R. (2002). Codon 12 and codon 13 mutations at the K-ras gene induce different soft tissue sarcoma types in nude mice. FASEB J *16*, 1642-1644.
Guha, U., Chaerkady, R., Marimuthu, A., Patterson, A.S., Kashyap, M.K., Harsha, H.C., Sato, M., Bader, J.S., Lash, A.E., Minna, J.D.*, et al.* (2008). Comparisons of tyrosine phosphorylated proteins in cells expressing lung cancer-specific alleles of EGFR and KRAS. Proc Natl Acad Sci U S A *105*, 14112-14117.
Guil, S., de La Iglesia, N., Fernandez-Larrea, J., Cifuentes, D., Ferrer, J.C., Guinovart, J.J., and Bach-Elias, M. (2003). Alternative splicing of the human proto-oncogene c-H-ras renders a new Ras family protein that trafficks to cytoplasm and nucleus. Cancer Res *63*, 5178-5187.
Gutierrez, L., Magee, A.I., Marshall, C.J., and Hancock, J.F. (1989). Post-translational processing of p21ras is two-step and involves carboxyl-methylation and carboxy-terminal proteolysis. EMBO J *8*, 1093-1098.

Gygi, S.P., Rist, B., Gerber, S.A., Turecek, F., Gelb, M.H., and Aebersold, R. (1999a). Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat Biotechnol *17*, 994-999.
Gygi, S.P., Rochon, Y., Franza, B.R., and Aebersold, R. (1999b). Correlation between protein and mRNA abundance in yeast. Mol Cell Biol *19*, 1720-1730.
Haddon, W.F., and McLafferty, F.W. (1968). Metastable ion characteristics. VII. Collision-induced metastables. Journal of the American Chemical Society *90*, 4745-4746.
Hager, J.W. (2001). Axial ejection in a multipole mass spectrometer (US Patent Office).
Hager, J.W. (2002). A new linear ion trap mass spectrometer. Rapid Communications in Mass Spectrometry *16*, 512-526.
Haigis, K.M., Kendall, K.R., Wang, Y., Cheung, A., Haigis, M.C., Glickman, J.N., Niwa-Kawakita, M., Sweet-Cordero, A., Sebolt-Leopold, J., Shannon, K.M.*, et al.* (2008). Differential effects of oncogenic K-Ras and N-Ras on proliferation, differentiation and tumor progression in the colon. Nat Genet *40*, 600-608.
Hall, A., Marshall, C.J., Spurr, N.K., and Weiss, R.A. (1983). Identification of transforming gene in two human sarcoma cell lines as a new member of the ras gene family located on chromosome 1. Nature *303*, 396-400.
Halvey, P.J., Ferrone, C.R., and Liebler, D.C. (2012). GeLC-MRM quantitation of mutant KRAS oncoprotein in complex biological samples. J Proteome Res *11*, 3908-3913.
Hamilton, M., and Wolfman, A. (1998). Ha-ras and N-ras regulate MAPK activity by distinct mechanisms in vivo. Oncogene *16*, 1417-1428.
Hammond, D.E., Hyde, R., Kratchmarova, I., Beynon, R.J., Blagoev, B., and Clague, M.J. (2010). Quantitative analysis of HGF and EGF-dependent phosphotyrosine signaling networks. J Proteome Res *9*, 2734-2742.
Hanahan, D., and Weinberg, R.A. (2000). The hallmarks of cancer. Cell *100*, 57-70.
Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. Cell *144*, 646-674.
Hancock, J.F. (2003). Ras proteins: different signals from different locations. Nat Rev Mol Cell Biol *4*, 373-384.
Hancock, J.F., and Parton, R.G. (2005). Ras plasma membrane signalling platforms. Biochem J *389*, 1-11.
Hancock, J.F., Paterson, H., and Marshall, C.J. (1990). A polybasic domain or palmitoylation is required in addition to the CAAX motif to localize p21ras to the plasma membrane. Cell *63*, 133-139.
Hanke, S., Besir, H., Oesterhelt, D., and Mann, M. (2008). Absolute SILAC for accurate quantitation of proteins in complex mixtures down to the attomole level. J Proteome Res *7*, 1118-1130.
Harding, A., and Hancock, J.F. (2008). Ras nanoclusters: Combining digital and analog signaling. Cell Cycle *7*, 127-134.
Harrison, A.G. (1998). The gas-phase basicities and proton affinities of amino acids and peptides. Mass Spectrom Rev *16*, 201-217.
Harvey, J.J. (1964). An Unidentified Virus Which Causes the Rapid Production of Tumours in Mice. Nature *204*, 1104-1105.
Hernandez-Valladares, M., Aran, V., and Prior, I.A. (2014). Quantitative proteomic analysis of compartmentalized signaling networks. Methods Enzymol *535*, 309-325.
Hoffert, J.D., Pisitkun, T., Saeed, F., Song, J.H., Chou, C.L., and Knepper, M.A. (2012). Dynamics of the G protein-coupled vasopressin V2 receptor signaling network revealed by quantitative phosphoproteomics. Mol Cell Proteomics *11*, M111 014613.

Hrycyna, C.A., Sapperstein, S.K., Clarke, S., and Michaelis, S. (1991). The Saccharomyces cerevisiae STE14 gene encodes a methyltransferase that mediates C-terminal methylation of a-factor and RAS proteins. EMBO J *10*, 1699-1709.
Hsu, J.L., Huang, S.Y., Chow, N.H., and Chen, S.H. (2003). Stable-isotope dimethyl labeling for quantitative proteomics. Anal Chem *75*, 6843-6852.
Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009a). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc *4*, 44-57.
Huang da, W., Sherman, B.T., Zheng, X., Yang, J., Imamichi, T., Stephens, R., and Lempicki, R.A. (2009b). Extracting biological meaning from large gene lists with DAVID. Curr Protoc Bioinformatics *Chapter 13*, Unit 13 11.
Huang, L., Hofer, F., Martin, G.S., and Kim, S.H. (1998). Structural basis for the interaction of Ras with RalGDS. Nat Struct Biol *5*, 422-426.
Hunter, T. (1991). Protein kinase classification. Methods Enzymol *200*, 3-37.
Hunter, T. (1998). The Croonian Lecture 1997. The phosphorylation of proteins on tyrosine: its role in cell growth and disease. Philos Trans R Soc Lond B Biol Sci *353*, 583-605.
Hunter, T., and Sefton, B.M. (1980). Transforming gene product of Rous sarcoma virus phosphorylates tyrosine. Proc Natl Acad Sci U S A *77*, 1311-1315.
Imamura, Y., Morikawa, T., Liao, X., Lochhead, P., Kuchiba, A., Yamauchi, M., Qian, Z.R., Nishihara, R., Meyerhardt, J.A., Haigis, K.M*., et al.* (2012). Specific mutations in KRAS codons 12 and 13, and patient prognosis in 1075 BRAF wild-type colorectal cancers. Clin Cancer Res *18*, 4753-4763.
Inder, K., Harding, A., Plowman, S.J., Philips, M.R., Parton, R.G., and Hancock, J.F. (2008). Activation of the MAPK module from different spatial locations generates distinct system outputs. Mol Biol Cell *19*, 4776-4784.
Ishihama, Y., Wei, F.Y., Aoshima, K., Sato, T., Kuromitsu, J., and Oda, Y. (2007). Enhancement of the efficiency of phosphoproteomic identification by removing phosphates after phosphopeptide enrichment. J Proteome Res *6*, 1139-1144.
Iwata, S., Lee, J.W., Okada, K., Lee, J.K., Iwata, M., Rasmussen, B., Link, T.A., Ramaswamy, S., and Jap, B.K. (1998). Complete structure of the 11-subunit bovine mitochondrial cytochrome bc1 complex. Science *281*, 64-71.
Janes, K.A., and Lauffenburger, D.A. (2013). Models of signalling networks - what cell biologists can gain from them and give to them. J Cell Sci *126*, 1913-1921.
Jaumot, M., Yan, J., Clyde-Smith, J., Sluimer, J., and Hancock, J.F. (2002). The linker domain of the Ha-Ras hypervariable region regulates interactions with exchange factors, Raf-1 and phosphoinositide 3-kinase. J Biol Chem *277*, 272-278.
Jeng, H.H., Taylor, L.J., and Bar-Sagi, D. (2012). Sos-mediated cross-activation of wild-type Ras by oncogenic Ras is essential for tumorigenesis. Nat Commun *3*, 1168.
Jennings, K.R. (1968). Collision-induced decompositions of aromatic molecular ions. International Journal of Mass Spectrometry and Ion Physics *1*, 227-235.
Jeno, P., Mini, T., Moes, S., Hintermann, E., and Horst, M. (1995). Internal sequences from proteins digested in polyacrylamide gels. Anal Biochem *224*, 75-82.
Jeong, M.H., Bae, J., Kim, W.H., Yoo, S.M., Kim, J.W., Song, P.I., and Choi, K.H. (2006). p19ras interacts with and activates p73 by involving the MDM2 protein. J Biol Chem *281*, 8707-8715.
John, J., Frech, M., and Wittinghofer, A. (1988). Biochemical properties of Ha-ras encoded p21 mutants and mechanism of the autophosphorylation reaction. J Biol Chem *263*, 11792-11799.
John, J., Schlichting, I., Schiltz, E., Rosch, P., and Wittinghofer, A. (1989). C-terminal truncation if p21H preserves crucial kinetic and structural properties. J Bio Chem *264*, 13086-13092.

Johnson, L., Greenbaum, D., Cichowski, K., Mercer, K., Murphy, E., Schmitt, E., Bronson, R.T., Umanoff, H., Edelmann, W., Kucherlapati, R.*, et al.* (1997). K-ras is an essential gene in the mouse with partial functional overlap with N-ras. Genes Dev *11*, 2468-2481.
Jones, M.K., and Jackson, J.H. (1998). Ras-GRF activates Ha-Ras, but not N-Ras or K-Ras 4B, protein in vivo. J Biol Chem *273*, 1782-1787.
Jonker, D.J., O'Callaghan, C.J., Karapetis, C.S., Zalcberg, J.R., Tu, D., Au, H.J., Berry, S.R., Krahn, M., Price, T., Simes, R.J.*, et al.* (2007). Cetuximab for the treatment of colorectal cancer. N Engl J Med *357*, 2040-2048.
Jorgensen, C., Sherman, A., Chen, G.I., Pasculescu, A., Poliakov, A., Hsiung, M., Larsen, B., Wilkinson, D.G., Linding, R., and Pawson, T. (2009). Cell-specific information processing in segregating populations of Eph receptor ephrin-expressing cells. Science *326*, 1502-1509.
Jura, N., Scotto-Lavino, E., Sobczyk, A., and Bar-Sagi, D. (2006). Differential modification of Ras proteins by ubiquitination. Mol Cell *21*, 679-687.
Karapetis, C.S., Khambata-Ford, S., Jonker, D.J., O'Callaghan, C.J., Tu, D., Tebbutt, N.C., Simes, R.J., Chalchal, H., Shapiro, J.D., Robitaille, S.*, et al.* (2008). K-ras mutations and benefit from cetuximab in advanced colorectal cancer. N Engl J Med *359*, 1757-1765.
Karnoub, A.E., and Weinberg, R.A. (2008). Ras oncogenes: split personalities. Nat Rev Mol Cell Biol *9*, 517-531.
Keller, J.W., Franklin, J.L., Graves-Deal, R., Friedman, D.B., Whitwell, C.W., and Coffey, R.J. (2007a). Oncogenic KRAS provides a uniquely powerful and variable oncogenic contribution among RAS family members in the colonic epithelium. J Cell Physiol *210*, 740-749.
Keller, J.W., Haigis, K.M., Franklin, J.L., Whitehead, R.H., Jacks, T., and Coffey, R.J. (2007b). Oncogenic K-RAS subverts the antiapoptotic role of N-RAS and alters modulation of the N-RAS:gelsolin complex. Oncogene *26*, 3051-3059.
Kersey, P.J., Duarte, J., Williams, A., Karavidopoulou, Y., Birney, E., and Apweiler, R. (2004). The International Protein Index: an integrated database for proteomics experiments. Proteomics *4*, 1985-1988.
Kim, E., Ambroziak, P., Otto, J.C., Taylor, B., Ashby, M., Shannon, K., Casey, P.J., and Young, S.G. (1999). Disruption of the mouse Rce1 gene results in defective Ras processing and mislocalization of Ras within cells. J Biol Chem *274*, 8383-8390.
Kingdon, K. (1923). A Method for the Neutralization of Electron Space Charge by Positive Ionization at Very Low Gas Pressures. Physical Review *21*, 408-418.
Kirsten, W.H., and Mayer, L.A. (1967). Morphologic responses to a murine erythroblastosis virus. J Natl Cancer Inst *39*, 311-335.
Klipp, E., and Liebermeister, W. (2006). Mathematical modeling of intracellular signaling pathways. BMC Neurosci *7 Suppl 1*, S10.
Knight, R.D. (1981). Storage of ions from laser-produced plasmas. Applied Physics Letters *38*, 221.
Koera, K., Nakamura, K., Nakao, K., Miyoshi, J., Toyoshima, K., Hatta, T., Otani, H., Aiba, A., and Katsuki, M. (1997). K-ras is essential for the development of the mouse embryo. Oncogene *15*, 1151-1159.
Kohlbacher, O., Reinert, K., Gropl, C., Lange, E., Pfeifer, N., Schulz-Trieglaff, O., and Sturm, M. (2007). TOPP--the OpenMS proteomics pipeline. Bioinformatics *23*, e191-197.
Koide, H., Satoh, T., Nakafuku, M., and Kaziro, Y. (1993). GTP-dependent association of Raf-1 with Ha-Ras: identification of Raf as a target downstream of Ras in mammalian cells. Proc Natl Acad Sci U S A *90*, 8683-8686.

Kranenburg, O., Verlaan, I., and Moolenaar, W.H. (2001). Regulating c-Ras function. cholesterol depletion affects caveolin association, GTP loading, and signaling. Curr Biol *11*, 1880-1884.
Krengel, U., Schlichting, I., Scherer, A., Schumann, R., Frech, M., John, J., Kabsch, W., Pai, E.F., and Wittinghofer, A. (1990). Three-dimensional structures of H-ras p21 mutants: molecular basis for their inability to function as signal switch molecules. Cell *62*, 539-548.
Krontiris, T.G., and Cooper, G.M. (1981). Transforming activity of human tumor DNAs. Proc Natl Acad Sci U S A *78*, 1181-1184.
Lampson, B.L., Pershing, N.L., Prinz, J.A., Lacsina, J.R., Marzluff, W.F., Nicchitta, C.V., MacAlpine, D.M., and Counter, C.M. (2013). Rare codons regulate KRas oncogenesis. Curr Biol *23*, 70-75.
Langbeheim, H., Shih, T.Y., and Scolnick, E.M. (1980). Identification of a normal vertebrate cell protein related to the p21 src of Harvey murine sarcoma virus. Virology *106*, 292-300.
Lange, V., Picotti, P., Domon, B., and Aebersold, R. (2008). Selected reaction monitoring for quantitative proteomics: a tutorial. Mol Syst Biol *4*, 222.
Larsen, M.R., Thingholm, T.E., Jensen, O.N., Roepstorff, P., and Jorgensen, T.J. (2005). Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. Mol Cell Proteomics *4*, 873-886.
Laude, A.J., and Prior, I.A. (2008). Palmitoylation and localisation of RAS isoforms are modulated by the hypervariable linker domain. J Cell Sci *121*, 421-427.
Le Hellard, S., Havik, B., Espeseth, T., Breilid, H., Lovlie, R., Luciano, M., Gow, A.J., Harris, S.E., Starr, J.M., Wibrand, K.*, et al.* (2009). Variants in doublecortin- and calmodulin kinase like 1, a gene up-regulated by BDNF, are associated with memory and general cognitive abilities. PLoS One *4*, e7534.
Lee, S.Y., Kim, J.W., Jeong, M.H., An, J.H., Jang, S.M., Song, K.H., and Choi, K.H. (2008). Microtubule-associated protein 1B light chain (MAP1B-LC1) negatively regulates the activity of tumor suppressor p53 in neuroblastoma cells. FEBS Lett *582*, 2826-2832.
Lehmann, W.D., Kruger, R., Salek, M., Hung, C.W., Wolschin, F., and Weckwerth, W. (2007). Neutral loss-based phosphopeptide recognition: a collection of caveats. J Proteome Res *6*, 2866-2873.
Lemeer, S., Pinkse, M.W.H., Mohammed, S., van Breukelen, B., den Hertog, J., Slijper, M., and Heck, A.J.R. (2008). Online Automated in VivoZebrafish Phosphoproteomics: From Large-Scale Analysis Down to a Single Embryo. Journal of Proteome Research *7*, 1555-1564.
Leon, J., Guerrero, I., and Pellicer, A. (1987). Differential expression of the ras gene family in mice. Mol Cell Biol *7*, 1535-1540.
Liao, P.C., Leykam, J., Andrews, P.C., Gage, D.A., and Allison, J. (1994). An approach to locate phosphorylation sites in a phosphoprotein: mass mapping by combining specific enzymatic degradation with matrix-assisted laser desorption/ionization mass spectrometry. Anal Biochem *219*, 9-20.
Lievre, A., Bachet, J.B., Boige, V., Cayre, A., Le Corre, D., Buc, E., Ychou, M., Bouche, O., Landi, B., Louvet, C.*, et al.* (2008). KRAS mutations as an independent prognostic factor in patients with advanced colorectal cancer treated with cetuximab. J Clin Oncol *26*, 374-379.
Lievre, A., Bachet, J.B., Le Corre, D., Boige, V., Landi, B., Emile, J.F., Cote, J.F., Tomasic, G., Penna, C., Ducreux, M.*, et al.* (2006). KRAS mutation status is predictive of response to cetuximab therapy in colorectal cancer. Cancer Res *66*, 3992-3995.
Lim, K.H., Ancrile, B.B., Kashatus, D.F., and Counter, C.M. (2008). Tumour maintenance is mediated by eNOS. Nature *452*, 646-649.

Linding, R., Jensen, L.J., Pasculescu, A., Olhovsky, M., Colwill, K., Bork, P., Yaffe, M.B., and Pawson, T. (2008). NetworKIN: a resource for exploring cellular phosphorylation networks. Nucleic Acids Res *36*, D695-699.
Little, A.S., Balmanno, K., Sale, M.J., Newman, S., Dry, J.R., Hampson, M., Edwards, P.A., Smith, P.D., and Cook, S.J. (2011). Amplification of the driving oncogene, KRAS or BRAF, underpins acquired resistance to MEK1/2 inhibitors in colorectal cancer cells. Sci Signal *4*, ra17.
Liu, H., Sadygov, R.G., and Yates, J.R., 3rd (2004). A model for random sampling and estimation of relative protein abundance in shotgun proteomics. Anal Chem *76*, 4193-4201.
Liu, L.K., Busch, K.L., and Cooks, R.G. (1981). Matrix-Assisted Secondary Ion Mass-Spectra of Biological Compounds. Analytical Chemistry *53*, 109-113.
Liu, W., Guan, M., Hu, T., Gu, X., and Lu, Y. (2011). Re-expression of AKAP12 inhibits progression and metastasis potential of colorectal carcinoma in vivo and in vitro. PLoS One *6*, e24015.
Locker, G.Y., Hamilton, S., Harris, J., Jessup, J.M., Kemeny, N., Macdonald, J.S., Somerfield, M.R., Hayes, D.F., and Bast, R.C. (2006). ASCO 2006 update of recommendations for the use of tumor markers in gastrointestinal cancer. Journal of Clinical Oncology *24*, 5313-5327.
Londry, F.A., and Hager, J.W. (2003). Mass selective axial ion ejection from a linear quadrupole ion trap. J Am Soc Mass Spectrom *14*, 1130-1147.
Lorenzo, P.S., Kung, J.W., Bottorff, D.A., Garfield, S.H., Stone, J.C., and Blumberg, P.M. (2001). Phorbol esters modulate the Ras exchange factor RasGRP3. Cancer Res *61*, 943-949.
Lowenthal, M.S., Liang, Y., Phinney, K.W., and Stein, S.E. (2014). Quantitative bottom-up proteomics depends on digestion conditions. Analytical Chemistry *86*, 551-558.
Luo, J., Emanuele, M.J., Li, D., Creighton, C.J., Schlabach, M.R., Westbrook, T.F., Wong, K.K., and Elledge, S.J. (2009). A genome-wide RNAi screen identifies multiple synthetic lethal interactions with the Ras oncogene. Cell *137*, 835-848.
Macek, B., Gnad, F., Soufi, B., Kumar, C., Olsen, J.V., Mijakovic, I., and Mann, M. (2008). Phosphoproteome analysis of E. coli reveals evolutionary conservation of bacterial Ser/Thr/Tyr phosphorylation. Mol Cell Proteomics *7*, 299-307.
MacLean, B., Tomazela, D.M., Shulman, N., Chambers, M., Finney, G.L., Frewen, B., Kern, R., Tabb, D.L., Liebler, D.C., and MacCoss, M.J. (2010). Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics *26*, 966-968.
Magee, A.I., Gutierrez, L., McKay, I.A., Marshall, C.J., and Hall, A. (1987). Dynamic fatty acylation of p21N-ras. EMBO J *6*, 3353-3357.
Makarov, A. (2000). Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis. Anal Chem *72*, 1156-1162.
Makarov, A., Denisov, E., Kholomeev, A., Balschun, W., Lange, O., Strupat, K., and Horning, S. (2006a). Performance evaluation of a hybrid linear ion trap/orbitrap mass spectrometer. *78*, 2113-2120.
Makarov, A., Denisov, E., Lange, O., and Horning, S. (2006b). Dynamic range of mass accuracy in LTQ Orbitrap hybrid mass spectrometer. J Am Soc Mass Spectrom *17*, 977-982.
Makarov, A., and Scigelova, M. (2010). Coupling liquid chromatography to Orbitrap mass spectrometry. J Chromatogr A *1217*, 3938-3945.
Malbon, C.C., Tao, J., Shumay, E., and Wang, H.Y. (2004). AKAP (A-kinase anchoring protein) domains: beads of structure-function on the necklace of G-protein signalling. Biochem Soc Trans *32*, 861-864.

Mallick, P., Schirle, M., Chen, S.S., Flory, M.R., Lee, H., Martin, D., Ranish, J., Raught, B., Schmitt, R., Werner, T.*, et al.* (2007). Computational prediction of proteotypic peptides for quantitative proteomics. Nat Biotechnol *25*, 125-131.
Malumbres, M., and Barbacid, M. (2003). RAS oncogenes: the first 30 years. Nat Rev Cancer *3*, 459-465.
Manning, G., Whyte, D.B., Martinez, R., Hunter, T., and Sudarsanam, S. (2002). The protein kinase complement of the human genome. Science *298*, 1912-1934.
Mao, C., Huang, Y.F., Yang, Z.Y., Zheng, D.Y., Chen, J.Z., and Tang, J.L. (2013). KRAS p.G13D mutation and codon 12 mutations are not created equal in predicting clinical outcomes of cetuximab in metastatic colorectal cancer: a systematic review and meta-analysis. Cancer *119*, 714-721.
Marais, R., Light, Y., Mason, C., Paterson, H., Olson, M.F., and Marshall, C.J. (1998). Requirement of Ras-GTP-Raf complexes for activation of Raf-1 by protein kinase C. Science *280*, 109-112.
Marais, R., Light, Y., Paterson, H.F., and Marshall, C.J. (1995). Ras recruits Raf-1 to the plasma membrane for activation by tyrosine phosphorylation. EMBO J *14*, 3136-3145.
March, R.E. (1997). An introduction to quadrupole ion trap mass spectrometry. J Mass Spectrom *32*, 351-369.
Marcotte, E.M. (2007). How do shotgun proteomics algorithms identify proteins? Nat Biotechnol *25*, 755-757.
Matallanas, D., Arozarena, I., Berciano, M.T., Aaronson, D.S., Pellicer, A., Lafarga, M., and Crespo, P. (2003). Differences on the inhibitory specificities of H-Ras, K-Ras, and N-Ras (N17) dominant negative mutants are related to their membrane microlocalization. J Biol Chem *278*, 4572-4581.
Matallanas, D., Sanz-Moreno, V., Arozarena, I., Calvo, F., Agudo-Ibanez, L., Santos, E., Berciano, M.T., and Crespo, P. (2006). Distinct utilization of effectors and biological outcomes resulting from site-specific Ras activation: Ras functions in lipid rafts and Golgi complex are dispensable for proliferation and transformation. Mol Cell Biol *26*, 100-116.
McGrath, J.P., Capon, D.J., Smith, D.H., Chen, E.Y., Seeburg, P.H., Goeddel, D.V., and Levinson, A.D. (1983). Structure and organization of the human Ki-ras proto-oncogene and a related processed pseudogene. Nature *304*, 501-506.
McLafferty, F.W., and Bryce, T.A. (1967). Metastable-ion characteristics: characterization of isomeric molecules. 1215-1217.
McLeod, H.L., and Church, R.D. (2003). Molecular predictors of prognosis and response to therapy in colorectal cancer. Cancer Chemother Biol Response Modif *21*, 791-801.
Messner, I., Cadeddu, G., Huckenbeck, W., Knowles, H.J., Gabbert, H.E., Baldus, S.E., and Schaefer, K.L. (2013). KRAS p.G13D mutations are associated with sensitivity to anti-EGFR antibody treatment in colorectal cancer cell lines. J Cancer Res Clin Oncol *139*, 201-209.
Mikesh, L.M., Ueberheide, B., Chi, A., Coon, J.J., Syka, J.E., Shabanowitz, J., and Hunt, D.F. (2006). The utility of ETD mass spectrometry in proteomic analysis. Biochim Biophys Acta *1764*, 1811-1822.
Milburn, M.V., Tong, L., deVos, A.M., Brunger, A., Yamaizumi, Z., Nishimura, S., and Kim, S.H. (1990). Molecular switch for signal transduction: structural differences between active and inactive forms of protooncogenic ras proteins. Science *247*, 939-945.
Millan, O., Ballester, A., Castrillo, A., Oliva, J.L., Traves, P.G., Rojas, J.M., and Bosca, L. (2003). H-Ras-specific activation of NF-kappaB protects NIH 3T3 cells against stimulus-dependent apoptosis. Oncogene *22*, 477-483.

Misale, S., Yaeger, R., Hobor, S., Scala, E., Janakiraman, M., Liska, D., Valtorta, E., Schiavo, R., Buscarino, M., Siravegna, G.*, et al.* (2012). Emergence of KRAS mutations and acquired resistance to anti-EGFR therapy in colorectal cancer. Nature *486*, 532-536.
Modest, D.P., Stintzing, S., Laubender, R.P., Neumann, J., Jung, A., Giessen, C., Haas, M., Aubele, P., Schulz, C., Boeck, S.*, et al.* (2011). Clinical characterization of patients with metastatic colorectal cancer depending on the KRAS status. Anticancer Drugs *22*, 913-918.
Monroe, M.E., Tolic, N., Jaitly, N., Shaw, J.L., Adkins, J.N., and Smith, R.D. (2007). VIPER: an advanced software package to support high-throughput LC-MS peptide identification. Bioinformatics *23*, 2021-2023.
Moodie, S.A., Willumsen, B.M., Weber, M.J., and Wolfman, A. (1993). Complexes of Ras.GTP with Raf-1 and mitogen-activated protein kinase kinase. Science *260*, 1658-1661.
Morandell, S., Stasyk, T., Skvortsov, S., Ascher, S., and Huber, L.A. (2008). Quantitative proteomics and phosphoproteomics reveal novel insights into complexity and dynamics of the EGFR signaling network. Proteomics *8*, 4383-4401.
Moroni, M., Veronese, S., Benvenuti, S., Marrapese, G., Sartore-Bianchi, A., Di Nicolantonio, F., Gambacorta, M., Siena, S., and Bardelli, A. (2005). Gene copy number for epidermal growth factor receptor (EGFR) and clinical response to antiEGFR treatment in colorectal cancer: a cohort study. The Lancet Oncology *6*, 279-286.
Muraoka, S., Shima, F., Araki, M., Inoue, T., Yoshimoto, A., Ijiri, Y., Seki, N., Tamura, A., Kumasaka, T., Yamamoto, M.*, et al.* (2012). Crystal structures of the state 1 conformations of the GTP-bound H-Ras protein and its oncogenic G12V and Q61L mutants. FEBS Lett *586*, 1715-1718.
Murray, M.J., Shilo, B.Z., Shih, C., Cowing, D., Hsu, H.W., and Weinberg, R.A. (1981). Three different human tumor cell lines contain different oncogenes. Cell *25*, 355-361.
Nagaraj, N., Wisniewski, J.R., Geiger, T., Cox, J., Kircher, M., Kelso, J., Paabo, S., and Mann, M. (2011). Deep proteome and transcriptome mapping of a human cancer cell line. Mol Syst Biol *7*, 548.
Nakanishi, Y., Seno, H., Fukuoka, A., Ueo, T., Yamaga, Y., Maruno, T., Nakanishi, N., Kanda, K., Komekado, H., Kawada, M.*, et al.* (2013). Dclk1 distinguishes between tumor and normal stem cells in the intestine. Nat Genet *45*, 98-103.
Nakano, H., Yamamoto, F., Neville, C., Evans, D., Mizuno, T., and Perucho, M. (1984). Isolation of transforming sequences of two human lung carcinomas: structural and functional analysis of the activated c-K-ras oncogenes. Proc Natl Acad Sci U S A *81*, 71-75.
Nappi, M., Weil, C., Cleven, C.D., Horn, L.A., Wollnik, H., and Cooks, R.G. (1997). Visual representations of simulated three-dimensional ion trajectories in an ion trap mass spectrometer. International Journal of Mass Spectrometry and Ion Processes *161*, 77-85.
Nassar, N., Horn, G., Herrmann, C., Block, C., Janknecht, R., and Wittinghofer, A. (1996). Ras/Rap effector specificity determined by charge reversal. Nat Struct Biol *3*, 723-729.
Nassar, N., Horn, G., Herrmann, C., Scherer, A., McCormick, F., and Wittinghofer, A. (1995). The 2.2 A crystal structure of the Ras-binding domain of the serine/threonine kinase c-Raf1 in complex with Rap1A and a GTP analogue. Nature *375*, 554-560.
Nesvizhskii, A.I., Vitek, O., and Aebersold, R. (2007). Analysis and validation of proteomic data generated by tandem mass spectrometry. Nat Methods *4*, 787-797.

Nold, M.J., Cerda, B.A., and Wesdemiotis, C. (1999). Proton affinities of the N- and C-terminal segments arising upon the dissociation of the amide bond in protonated peptides. J Am Soc Mass Spectrom *10*, 1-8.
Oda, Y., Huang, K., Cross, F.R., Cowburn, D., and Chait, B.T. (1999). Accurate quantitation of protein expression and site-specific phosphorylation. Proc Natl Acad Sci U S A *96*, 6591-6596.
Ohba, Y., Mochizuki, N., Yamashita, S., Chan, A.M., Schrader, J.W., Hattori, S., Nagashima, K., and Matsuda, M. (2000). Regulatory proteins of R-Ras, TC21/R-Ras2, and M-Ras/R-Ras3. J Biol Chem *275*, 20020-20026.
Olsen, J.V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P., and Mann, M. (2006). Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. Cell *127*, 635-648.
Olsen, J.V., and Macek, B. (2009). High accuracy mass spectrometry in large-scale analysis of protein phosphorylation. Methods Mol Biol *492*, 131-142.
Olsen, J.V., Macek, B., Lange, O., Makarov, A., Horning, S., and Mann, M. (2007). Higher-energy C-trap dissociation for peptide modification analysis. Nat Methods *4*, 709-712.
Olsen, J.V., and Mann, M. (2004). Improved peptide identification in proteomics by two consecutive stages of mass spectrometric fragmentation. Proc Natl Acad Sci U S A *101*, 13417-13422.
Olsen, J.V., Ong, S.E., and Mann, M. (2004). Trypsin cleaves exclusively C-terminal to arginine and lysine residues. Mol Cell Proteomics *3*, 608-614.
Omerovic, J., Hammond, D.E., Clague, M.J., and Prior, I.A. (2008). Ras isoform abundance and signalling in human cancer cell lines. Oncogene *27*, 2754-2762.
Omerovic, J., Laude, A.J., and Prior, I.A. (2007). Ras proteins: paradigms for compartmentalised and isoform-specific signalling. Cell Mol Life Sci *64*, 2575-2589.
Omerovic, J., and Prior, I.A. (2009). Compartmentalized signalling: Ras proteins and signalling nanoclusters. FEBS J *276*, 1817-1825.
Ong, S.E., Blagoev, B., Kratchmarova, I., Kristensen, D.B., Steen, H., Pandey, A., and Mann, M. (2002). Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics *1*, 376-386.
Onken, B., Wiener, H., Philips, M.R., and Chang, E.C. (2006). Compartmentalized signaling of Ras in fission yeast. Proc Natl Acad Sci U S A *103*, 9045-9050.
Pace, C.N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995). How to measure and predict the molar absorption coefficient of a protein. Protein Sci *4*, 2411-2423.
Pacold, M.E., Suire, S., Perisic, O., Lara-Gonzalez, S., Davis, C.T., Walker, E.H., Hawkins, P.T., Stephens, L., Eccleston, J.F., and Williams, R.L. (2000). Crystal structure and functional analysis of Ras binding to its effector phosphoinositide 3-kinase gamma. Cell *103*, 931-943.
Pai, E.F., Kabsch, W., Krengel, U., Holmes, K.C., John, J., and Wittinghofer, A. (1989). Structure of the guanine-nucleotide-binding domain of the Ha-ras oncogene product p21 in the triphosphate conformation. Nature *341*, 209-214.
Pai, E.F., Krengel, U., Petsko, G.A., Goody, R.S., Kabsch, W., and Wittinghofer, A. (1990). Refined crystal structure of the triphosphate conformation of H-ras p21 at 1.35 A resolution: implications for the mechanism of GTP hydrolysis. EMBO J *9*, 2351-2359.
Paizs, B., and Suhai, S. (2001). Theoretical study of the main fragmentation pathways for protonated glycylglycine. Rapid Commun Mass Spectrom *15*, 651-663.
Paizs, B., and Suhai, S. (2004). Towards understanding the tandem mass spectra of protonated oligopeptides. 1: mechanism of amide bond cleavage. J Am Soc Mass Spectrom *15*, 103-113.

Paizs, B., and Suhai, S. (2005). Fragmentation pathways of protonated peptides. Mass Spectrom Rev *24*, 508-548.
Palumbo, A.M., and Reid, G.E. (2008). Evaluation of gas-phase rearrangement and competing fragmentation reactions on protein phosphorylation site assignment using collision induced dissociation-MS/MS and MS3. Anal Chem *80*, 9735-9747.
Palumbo, A.M., Tepe, J.J., and Reid, G.E. (2008). Mechanistic insights into the multistage gas-phase fragmentation behavior of phosphoserine- and phosphothreonine-containing peptides. J Proteome Res *7*, 771-779.
Parada, L.F., Tabin, C.J., Shih, C., and Weinberg, R.A. (1982). Human EJ bladder carcinoma oncogene is homologue of Harvey sarcoma virus ras gene. Nature *297*, 474-478.
Parajuli, B., Georgiadis, T.M., Fishel, M.L., and Hurley, T.D. (2014). Development of selective inhibitors for human aldehyde dehydrogenase 3A1 (ALDH3A1) for the enhancement of cyclophosphamide cytotoxicity. Chembiochem *15*, 701-712.
Park, S.K., Venable, J.D., Xu, T., and Yates, J.R., 3rd (2008). A quantitative analysis software tool for mass spectrometry-based proteomics. Nat Methods *5*, 319-322.
Parks, W.P., and Scolnick, E.M. (1977). In vitro translation of Harvey murine sarcoma virus RNA. Journal of virology *22*, 711-719.
Parsons, B.L., and Myers, M.B. (2013). Personalized cancer treatment and the myth of KRAS wild-type colon tumors. Discov Med *15*, 259-267.
Parton, R.G., and del Pozo, M.A. (2013). Caveolae as plasma membrane sensors, protectors and organizers. Nat Rev Mol Cell Biol *14*, 98-112.
Paša-Tolić, L., Jensen, P.K., Anderson, G.A., Lipton, M.S., Peden, K.K., Martinović, S., Tolić, N., Bruce, J.E., and Smith, R.D. (1999). High Throughput Proteome-Wide Precision Measurements of Protein Expression Using Mass Spectrometry. Journal of the American Chemical Society *121*, 7949-7950.
Paul, W., Reinhard, H.P., and Vonzahn, U. (1958). Das Elektrische Massenfilter Als Massenspektrometer Und Isotopentrenner. Zeitschrift Fur Physik *152*, 143-182.
Paul, W., and Steinwedel, H. (1953). Ein neues massenspektrometer ohne magnetfeld. *8*, 448.
Perkins, D.N., Pappin, D.J., Creasy, D.M., and Cottrell, J.S. (1999). Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis *20*, 3551-3567.
Perucho, M., Goldfarb, M., Shimizu, K., Lama, C., Fogh, J., and Wigler, M. (1981). Human-tumor-derived cell lines contain common and different transforming genes. Cell *27*, 467-476.
Philips, M.R. (2005). Compartmentalized signalling of Ras. Biochem Soc Trans *33*, 657-661.
Pinkse, M.W., Uitto, P.M., Hilhorst, M.J., Ooms, B., and Heck, A.J. (2004). Selective isolation at the femtomole level of phosphopeptides from proteolytic digests using 2D-NanoLC-ESI-MS/MS and titanium oxide precolumns. Anal Chem *76*, 3935-3943.
Piszkiewicz, D., Landon, M., and Smith, E.L. (1970). Anomalous cleavage of aspartyl-proline peptide bonds during amino acid sequence determinations. Biochem Biophys Res Commun *40*, 1173-1178.
Plowman, S.J., Berry, R.L., Bader, S.A., Luo, F., Arends, M.J., Harrison, D.J., Hooper, M.L., and Patek, C.E. (2006). K-ras 4A and 4B are co-expressed widely in human tissues, and their ratio is altered in sporadic colorectal cancer. J Exp Clin Cancer Res *25*, 259-267.
Plowman, S.J., Muncke, C., Parton, R.G., and Hancock, J.F. (2005). H-ras, K-ras, and inner plasma membrane raft proteins operate in nanoclusters with differential dependence on the actin cytoskeleton. Proc Natl Acad Sci U S A *102*, 15500-15505.
Plowman, S.J., Williamson, D.J., O'Sullivan, M.J., Doig, J., Ritchie, A.M., Harrison, D.J., Melton, D.W., Arends, M.J., Hooper, M.L., and Patek, C.E. (2003). While K-ras

is essential for mouse development, expression of the K-ras 4A splice variant is dispensable. Mol Cell Biol *23*, 9245-9250.
Polakis, P., and McCormick, F. (1993). Structural requirements for the interaction of p21ras with GAP, exchange factors, and its biological effector target. J Biol Chem *268*, 9157-9160.
Ponting, C.P., and Benjamin, D.R. (1996). A novel family of Ras-binding domains. Trends Biochem Sci *21*, 422-425.
Potenza, N., Vecchione, C., Notte, A., De Rienzo, A., Rosica, A., Bauer, L., Affuso, A., De Felice, M., Russo, T., Poulet, R.*, et al.* (2005). Replacement of K-Ras with H-Ras supports normal embryonic development despite inducing cardiovascular pathology in adult mice. EMBO Rep *6*, 432-437.
Prahallad, A., Sun, C., Huang, S., Di Nicolantonio, F., Salazar, R., Zecchin, D., Beijersbergen, R.L., Bardelli, A., and Bernards, R. (2012). Unresponsiveness of colon cancer to BRAF(V600E) inhibition through feedback activation of EGFR. Nature *483*, 100-103.
Prior, I.A., and Hancock, J.F. (2012). Ras trafficking, localization and compartmentalized signalling. Semin Cell Dev Biol *23*, 145-153.
Prior, I.A., Harding, A., Yan, J., Sluimer, J., Parton, R.G., and Hancock, J.F. (2001). GTP-dependent segregation of H-ras from lipid rafts is required for biological activity. Nature Cell Biology *3*, 368-375.
Prior, I.A., Lewis, P.D., and Mattos, C. (2012). A comprehensive survey of Ras mutations in cancer. Cancer Res *72*, 2457-2467.
Prior, I.A., Muncke, C., Parton, R.G., and Hancock, J.F. (2003). Direct visualization of Ras proteins in spatially distinct cell surface microdomains. J Cell Biol *160*, 165-170.
Prive, G.G., Milburn, M.V., Tong, L., de Vos, A.M., Yamaizumi, Z., Nishimura, S., and Kim, S.H. (1992). X-ray crystal structures of transforming p21 ras mutants suggest a transition-state stabilization mechanism for GTP hydrolysis. Proc Natl Acad Sci U S A *89*, 3649-3653.
Qin, J., and Chait, B.T. (1995). Preferential Fragmentation of Protonated Gas-Phase Peptide Ions Adjacent to Acidic Amino Acid Residues. Journal of the American Chemical Society *117*, 5411-5412.
Rappsilber, J., Mann, M., and Ishihama, Y. (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nature Protocols *2*, 1896-1906.
Reddy, E.P., Reynolds, R.K., Santos, E., and Barbacid, M. (1982). A point mutation is responsible for the acquisition of transforming properties by the T24 human bladder carcinoma oncogene. Nature *300*, 149-152.
Reiss, Y., Goldstein, J.L., Seabra, M.C., Casey, P.J., and Brown, M.S. (1990). Inhibition of purified p21ras farnesyl:protein transferase by Cys-AAX tetrapeptides. Cell *62*, 81-88.
Ren, J., Li, G., Ge, J., Li, X., and Zhao, Y. (2012). Is K-ras gene mutation a prognostic factor for colorectal cancer: a systematic review and meta-analysis. Dis Colon Rectum *55*, 913-923.
Rietschel, B., Arrey, T.N., Meyer, B., Bornemann, S., Schuerken, M., Karas, M., and Poetsch, A. (2009). Elastase digests: new ammunition for shotgun membrane proteomics. Mol Cell Proteomics *8*, 1029-1043.
Rigbolt, K.T., and Blagoev, B. (2012). Quantitative phosphoproteomics to characterize signaling networks. Semin Cell Dev Biol *23*, 863-871.
Rigbolt, K.T., Prokhorova, T.A., Akimov, V., Henningsen, J., Johansen, P.T., Kratchmarova, I., Kassem, M., Mann, M., Olsen, J.V., and Blagoev, B. (2011a). System-wide temporal characterization of the proteome and phosphoproteome of human embryonic stem cell differentiation. Sci Signal *4*, rs3.

Rigbolt, K.T., Vanselow, J.T., and Blagoev, B. (2011b). GProX, a user-friendly platform for bioinformatics analysis and visualization of quantitative proteomics data. Mol Cell Proteomics *10*, O110 007450.
Rocks, O., Peyker, A., Kahms, M., Verveer, P.J., Koerner, C., Lumbierres, M., Kuhlmann, J., Waldmann, H., Wittinghofer, A., and Bastiaens, P.I. (2005). An acylation cycle regulates localization and activity of palmitoylated Ras isoforms. Science *307*, 1746-1752.
Rodriguez-Viciana, P., Sabatier, C., and McCormick, F. (2004). Signaling specificity by Ras family GTPases is determined by the full spectrum of effectors they regulate. Mol Cell Biol *24*, 4943-4954.
Rogers, L.D., Brown, N.F., Fang, Y., Pelech, S., and Foster, L.J. (2011). Phosphoproteomic analysis of Salmonella-infected cells identifies key kinase regulators and SopB-dependent host phosphorylation events. Sci Signal *4*, rs9.
Roskoski, R., Jr. (2012). ERK1/2 MAP kinases: structure, function, and regulation. Pharmacol Res *66*, 105-143.
Ross, P.L., Huang, Y.N., Marchese, J.N., Williamson, B., Parker, K., Hattan, S., Khainovski, N., Pillai, S., Dey, S., Daniels, S.*, et al.* (2004). Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. Mol Cell Proteomics *3*, 1154-1169.
Rous, P. (1910). A Transmissible Avian Neoplasm. (Sarcoma of the Common Fowl.). J Exp Med *12*, 696-705.
Roy, S., Plowman, S., Rotblat, B., Prior, I.A., Muncke, C., Grainger, S., Parton, R.G., Henis, Y.I., Kloog, Y., and Hancock, J.F. (2005). Individual palmitoyl residues serve distinct roles in H-ras trafficking, microlocalization, and signaling. Mol Cell Biol *25*, 6722-6733.
Roy, S., Wyse, B., and Hancock, J.F. (2002). H-Ras signaling and K-Ras signaling are differentially dependent on endocytosis. Mol Cell Biol *22*, 5128-5140.
Roy, U.K.B., Henkhaus, R.S., Loupakis, F., Cremolini, C., Gerner, E.W., and Ignatenko, N.A. (2013). Caveolin-1 is a novel regulator of K-RAS-dependent migration in colon carcinogenesis. International Journal of Cancer *133*, 43-57.
Ruppen-Canas, I., Lopez-Casas, P.P., Garcia, F., Ximenez-Embun, P., Munoz, M., Morelli, M.P., Real, F.X., Serna, A., Hidalgo, M., and Ashman, K. (2012). An improved quantitative mass spectrometry analysis of tumor specific mutant proteins at high sensitivity. Proteomics *12*, 1319-1327.
Rush, J., Moritz, A., Lee, K.A., Guo, A., Goss, V.L., Spek, E.J., Zhang, H., Zha, X.M., Polakiewicz, R.D., and Comb, M.J. (2005). Immunoaffinity profiling of tyrosine phosphorylation in cancer cells. Nat Biotechnol *23*, 94-101.
Saif, M.W., and Chu, E. (2010). Biology of colorectal cancer. Cancer J *16*, 196-201.
Saleem, R.A., Rogers, R.S., Ratushny, A.V., Dilworth, D.J., Shannon, P.T., Shteynberg, D., Wan, Y., Moritz, R.L., Nesvizhskii, A.I., Rachubinski, R.A.*, et al.* (2010). Integrated phosphoproteomics analysis of a signaling network governing nutrient response and peroxisome induction. Mol Cell Proteomics *9*, 2076-2088.
Samowitz, W.S., Curtin, K., Schaffer, D., Robertson, M., Leppert, M., and Slattery, M.L. (2000). Relationship of Ki-ras mutations in colon cancers to tumor location, stage, and survival: a population-based study. Cancer Epidemiol Biomarkers Prev *9*, 1193-1197.
Santos, E., Martin-Zanca, D., Reddy, E., Pierotti, M., Della Porta, G., and Barbacid, M. (1984). Malignant activation of a K-ras oncogene in lung carcinoma but not in normal tissue of the same patient. Science *223*, 661-664.
Santos, E., Tronick, S.R., Aaronson, S.A., Pulciani, S., and Barbacid, M. (1982). T24 human bladder carcinoma oncogene is an activated form of the normal human homologue of BALB- and Harvey-MSV transforming genes. Nature *298*, 343-347.

Saraste, M., Sibbald, P.R., and Wittinghofer, A. (1990). The P-loop--a common motif in ATP- and GTP-binding proteins. Trends Biochem Sci *15*, 430-434.
Sarthy, A.V., Morgan-Lappe, S.E., Zakula, D., Vernetti, L., Schurdak, M., Packer, J.C., Anderson, M.G., Shirasawa, S., Sasazuki, T., and Fesik, S.W. (2007). Survivin depletion preferentially reduces the survival of activated K-Ras-transformed cells. Mol Cancer Ther *6*, 269-276.
Sasaki, A.T., Carracedo, A., Locasale, J.W., Anastasiou, D., Takeuchi, K., Kahoud, E.R., Haviv, S., Asara, J.M., Pandolfi, P.P., and Cantley, L.C. (2011). Ubiquitination of K-Ras enhances activation and facilitates binding to select downstream effectors. Sci Signal *4*, ra13.
Savitski, M.M., Lemeer, S., Boesche, M., Lang, M., Mathieson, T., Bantscheff, M., and Kuster, B. (2011). Confident phosphorylation site localization using the Mascot Delta Score. Mol Cell Proteomics *10*, M110 003830.
Scheffzek, K., Ahmadian, M.R., Kabsch, W., Wiesmuller, L., Lautwein, A., Schmitz, F., and Wittinghofer, A. (1997). The Ras-RasGAP complex: structural basis for GTPase activation and its loss in oncogenic Ras mutants. Science *277*, 333-338.
Scheidig, A.J., Burmester, C., and Goody, R.S. (1999). The pre-hydrolysis state of p21(ras) in complex with GTP: new insights into the role of water molecules in the GTP hydrolysis reaction of ras-like proteins. Structure *7*, 1311-1324.
Scholl, C., Frohling, S., Dunn, I.F., Schinzel, A.C., Barbie, D.A., Kim, S.Y., Silver, S.J., Tamayo, P., Wadlow, R.C., Ramaswamy, S*., et al.* (2009). Synthetic Lethal Interaction between Oncogenic KRAS Dependency and STK33 Suppression in Human Cancer Cells. Cell *137*, 821-834.
Schreiber, T.B., Mausbacher, N., Soroka, J., Wandinger, S.K., Buchner, J., and Daub, H. (2012). Global analysis of phosphoproteome regulation by the Ser/Thr phosphatase Ppt1 in Saccharomyces cerevisiae. J Proteome Res *11*, 2397-2408.
Schroeder, M.J., Shabanowitz, J., Schwartz, J.C., Hunt, D.F., and Coon, J.J. (2004). A neutral loss activation method for improved phosphopeptide sequence analysis by quadrupole ion trap mass spectrometry. Anal Chem *76*, 3590-3598.
Schwanhausser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. Nature *473*, 337-342.
Schwartz, B.L., and Bursey, M.M. (1992). Some proline substituent effects in the tandem mass spectrum of protonated pentaalanine. Biol Mass Spectrom *21*, 92-96.
Schwartz, D., and Gygi, S.P. (2005). An iterative statistical approach to the identification of protein phosphorylation motifs from large-scale data sets. Nat Biotechnol *23*, 1391-1398.
Schwartz, J.C., Senko, M.W., and Syka, J.E. (2002). A two-dimensional quadrupole ion trap mass spectrometer. J Am Soc Mass Spectrom *13*, 659-669.
Schweins, T., Geyer, M., Scheffzek, K., Warshel, A., Kalbitzer, H.R., and Wittinghofer, A. (1995). Substrate-assisted catalysis as a mechanism for GTP hydrolysis of p21ras and other GTP-binding proteins. Nat Struc Biol *2*, 36-44.
Scolnick, E.M., Papageorge, A.G., and Shih, T.Y. (1979). Guanine nucleotide-binding activity as an assay for src protein of rat-derived murine sarcoma viruses. Proc Natl Acad Sci U S A *76*, 5355-5359.
Scolnick, E.M., and Parks, W.P. (1974). Harvey sarcoma virus: a second murine type C sarcoma virus with rat genetic information. J Virol *13*, 1211-1219.
Scolnick, E.M., Rands, E., Williams, D., and Parks, W.P. (1973). Studies on the nucleic acid sequences of Kirsten sarcoma virus: a model for formation of a mammalian RNA-containing sarcoma virus. J Virol *12*, 458-463.
Seeburg, P.H., Colby, W.W., Capon, D.J., Goeddel, D.V., and Levinson, A.D. (1984). Biological properties of human c-Ha-ras1 genes mutated at codon 12. Nature *312*, 71-75.

Shen, Y.F., Zhao, R., Berger, S.J., Anderson, G.A., Rodriguez, N., and Smith, R.D. (2002). High-efficiency nanoscale liquid chromatography coupled on-line with mass spectrometry using nanoelectrospray ionization for proteomics. Anal Chem *74*, 4235-4249.
Sherwood, C.A., Eastham, A., Lee, L.W., Risler, J., Mirzaei, H., Falkner, J.A., and Martin, D.B. (2009). Rapid optimization of MRM-MS instrument parameters by subtle alteration of precursor and product m/z targets. J Proteome Res *8*, 3746-3751.
Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996). Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. Anal Chem *68*, 850-858.
Shi, Y., Xiang, R., Horvath, C., and Wilkins, J.A. (2004). The role of liquid chromatography in proteomics. J Chromatogr A *1053*, 27-36.
Shih, C., Padhy, L.C., Murray, M., and Weinberg, R.A. (1981). Transforming genes of carcinomas and neuroblastomas introduced into mouse fibroblasts. Nature *290*, 261-264.
Shih, C., Shilo, B.Z., Goldfarb, M.P., Dannenberg, A., and Weinberg, R.A. (1979a). Passage of phenotypes of chemically transformed cells via transfection of DNA and chromatin. Proc Natl Acad Sci U S A *76*, 5714-5718.
Shih, C., and Weinberg, R.A. (1982). Isolation of a transforming sequence from a human bladder carcinoma cell line. Cell *29*, 161-169.
Shih, T.Y., Weeks, M.O., Young, H.A., and Scholnick, E.M. (1979b). Identification of a sarcoma virus-coded phosphoprotein in nonproducer cells transformed by Kirsten or Harvey murine sarcoma virus. Virology *96*, 64-79.
Shima, F., Ijiri, Y., Muraoka, S., Liao, J., Ye, M., Araki, M., Matsumoto, K., Yamamoto, N., Sugimoto, T., Yoshikawa, Y.*, et al.* (2010). Structural basis for conformational dynamics of GTP-bound Ras protein. J Biol Chem *285*, 22696-22705.
Shimizu, K., Goldfarb, M., Perucho, M., and Wigler, M. (1983). Isolation and preliminary characterization of the transforming gene of a human neuroblastoma cell line. Proc Natl Acad Sci U S A *80*, 383-387.
Siegel, R., DeSantis, C., Virgo, K., Stein, K., Mariotto, A., Smith, T., Cooper, D., Gansler, T., Lerro, C., Fedewa, S.*, et al.* (2012). Cancer treatment and survivorship statistics, 2012. CA Cancer J Clin *62*, 220-241.
Sigal, I.S., Gibbs, J.B., D'Alonzo, J.S., and Scolnick, E.M. (1986). Identification of effector residues and a neutralizing epitope of Ha-ras-encoded p21. Proc Natl Acad Sci U S A *83*, 4725-4729.
Silvius, J.R., Bhagatji, P., Leventis, R., and Terrone, D. (2006). K-ras4B and prenylated proteins lacking "second signals" associate dynamically with cellular membranes. Mol Biol Cell *17*, 192-202.
Singh, A., Sweeney, M.F., Yu, M., Burger, A., Greninger, P., Benes, C., Haber, D.A., and Settleman, J. (2012). TAK1 inhibition promotes apoptosis in KRAS-dependent colon cancers. Cell *148*, 639-650.
Singh, S., Springer, M., Steen, J., Kirschner, M.W., and Steen, H. (2009). FLEXIQuant: a novel tool for the absolute quantification of proteins, and the simultaneous identification and quantification of potentially modified peptides. J Proteome Res *8*, 2201-2210.
Sloan, S.R., Newcomb, E.W., and Pellicer, A. (1990). Neutron Radiation Can Activate K-Ras Via a Point Mutation in Codon 146 and Induces a Different Spectrum of Ras Mutations Than Does Gamma-Radiation. Molecular and Cellular Biology *10*, 405-408.
Smith, G., Bounds, R., Wolf, H., Steele, R.J., Carey, F.A., and Wolf, C.R. (2010). Activating K-Ras mutations outwith 'hotspot' codons in sporadic colorectal tumours - implications for personalised cancer medicine. Br J Cancer *102*, 693-703.

Smotrys, J.E., and Linder, M.E. (2004). Palmitoylation of intracellular signaling proteins: regulation and function. Annu Rev Biochem *73*, 559-587.
Spoerner, M., Herrmann, C., Vetter, I.R., Kalbitzer, H.R., and Wittinghofer, A. (2001). Dynamic properties of the Ras switch I region and its importance for binding to effectors. Proc Natl Acad Sci U S A *98*, 4944-4949.
Spoerner, M., Wittinghofer, A., and Kalbitzer, H.R. (2004). Perturbation of the conformational equilibria in Ras by selective mutations as studied by 31P NMR spectroscopy. FEBS Lett *578*, 305-310.
Stahl-Zeng, J., Lange, V., Ossola, R., Eckhardt, K., Krek, W., Aebersold, R., and Domon, B. (2007). High sensitivity detection of plasma proteins by multiple reaction monitoring of N-glycosites. Mol Cell Proteomics *6*, 1809-1817.
Steckel, M., Molina-Arcas, M., Weigelt, B., Marani, M., Warne, P.H., Kuznetsov, H., Kelly, G., Saunders, B., Howell, M., Downward, J.*, et al.* (2012). Determination of synthetic lethal interactions in KRAS oncogene-dependent cancer cells reveals novel therapeutic targeting strategies. Cell Res *22*, 1227-1245.
Steen, H., Jebanathirajah, J.A., Rush, J., Morrice, N., and Kirschner, M.W. (2006). Phosphorylation analysis by mass spectrometry: myths, facts, and the consequences for qualitative and quantitative measurements. Mol Cell Proteomics *5*, 172-181.
Steen, H., Kuster, B., Fernandez, M., Pandey, A., and Mann, M. (2001). Detection of tyrosine phosphorylated peptides by precursor ion scanning quadrupole TOF mass spectrometry in positive ion mode. Anal Chem *73*, 1440-1448.
Steen, H., and Mann, M. (2004). The ABC's (and XYZ's) of peptide sequencing. Nat Rev Mol Cell Biol *5*, 699-711.
Steinwedel, H., and Paul, W. (1960). Apparatus for separating charged particles of different specific charges. US Patent Office, 2,939,952.
Su, B., Bu, Y., Engelberg, D., and Gelman, I.H. (2010). SSeCKS/Gravin/AKAP12 inhibits cancer cell invasiveness and chemotaxis by suppressing a protein kinase C-Raf/MEK/ERK pathway. J Biol Chem *285*, 4578-4586.
Sudhir, P.R., Hsu, C.L., Wang, M.J., Wang, Y.T., Chen, Y.J., Sung, T.Y., Hsu, W.L., Yang, U.C., and Chen, J.Y. (2011). Phosphoproteomics Identifies Oncogenic Ras Signaling Targets and Their Involvement in Lung Adenocarcinomas. Plos One *6*.
Sugiyama, N., Masuda, T., Shinoda, K., Nakamura, A., Tomita, M., and Ishihama, Y. (2007). Phosphopeptide enrichment by aliphatic hydroxy acid-modified metal oxide chromatography for nano-LC-MS/MS in proteomics applications. Mol Cell Proteomics *6*, 1103-1109.
Sung, P.J., Rodrigues, A.B., Kleinberger, A., Quatela, S., Bach, E.A., and Philips, M.R. (2010). Cytosolic Ras supports eye development in Drosophila. Mol Cell Biol *30*, 5649-5657.
Swarthout, J.T., Lobo, S., Farh, L., Croke, M.R., Greentree, W.K., Deschenes, R.J., and Linder, M.E. (2005). DHHC9 and GCP16 constitute a human protein fatty acyltransferase with specificity for H- and N-Ras. J Biol Chem *280*, 31141-31148.
Sweet-Cordero, A., Mukherjee, S., Subramanian, A., You, H., Roix, J.J., Ladd-Acosta, C., Mesirov, J., Golub, T.R., and Jacks, T. (2005). An oncogenic KRAS2 expression signature identified by cross-species gene-expression analysis. Nat Genet *37*, 48-55.
Syka, J.E., Coon, J.J., Schroeder, M.J., Shabanowitz, J., and Hunt, D.F. (2004). Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. Proc Natl Acad Sci U S A *101*, 9528-9533.
Tabin, C.J., Bradley, S.M., Bargmann, C.I., Weinberg, R.A., Papageorge, A.G., Scolnick, E.M., Dhar, R., Lowy, D.R., and Chang, E.H. (1982). Mechanism of activation of a human oncogene. Nature *300*, 143-149.

Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida, Y., Yoshida, T., and Matsuo, T. (1988). Protein and polymer analyses up tom/z 100 000 by laser ionization time-of-flight mass spectrometry. Rapid Communications in Mass Spectrometry *2*, 151-153.
Taparowsky, E., Suard, Y., Fasano, O., Shimizu, K., Goldfarb, M., and Wigler, M. (1982). Activation of the T24 bladder carcinoma transforming gene is linked to a single amino acid change. Nature *300*, 762-765.
Taylor, G. (1964). Disintegration of Water Drops in an Electric Field. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences *280*, 383-397.
Tchernitsa, O.I., Sers, C., Zuber, J., Hinzmann, B., Grips, M., Schramme, A., Lund, P., Schwendel, A., Rosenthal, A., and Schafer, R. (2004). Transcriptional basis of KRAS oncogene-mediated cellular transformation in ovarian epithelial cells. Oncogene *23*, 4536-4555.
Tejpar, S., Celik, I., Schlichting, M., Sartorius, U., Bokemeyer, C., and Van Cutsem, E. (2012). Association of KRAS G13D tumor mutations with outcome in patients with metastatic colorectal cancer treated with first-line chemotherapy with or without cetuximab. J Clin Oncol *30*, 3570-3577.
Thermo (2008). LTQ Orbitrap XL Hardware Manual, Revision B - 1225830 edn.
Thissen, J.A., Gross, J.M., Subramanian, K., Meyer, T., and Casey, P.J. (1997). Prenylation-dependent association of Ki-Ras with microtubules. Evidence for a role in subcellular trafficking. J Biol Chem *272*, 30362-30370.
Tholey, A., Reed, J., and Lehmann, W.D. (1999). Electrospray tandem mass spectrometric studies of phosphopeptides and phosphopeptide analogues. J Mass Spectrom *34*, 117-123.
Thompson, A., Schafer, J., Kuhn, K., Kienle, S., Schwarz, J., Schmidt, G., Neumann, T., Johnstone, R., Mohammed, A.K., and Hamon, C. (2003). Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. Anal Chem *75*, 1895-1904.
Tian, T., Harding, A., Inder, K., Plowman, S., Parton, R.G., and Hancock, J.F. (2007). Plasma membrane nanoswitches generate high-fidelity Ras signal transduction. Nat Cell Biol *9*, 905-914.
To, M.D., Rosario, R.D., Westcott, P.M., Banta, K.L., and Balmain, A. (2013). Interactions between wild-type and mutant Ras genes in lung and skin carcinogenesis. Oncogene *32*, 4028-4033.
Trahey, M., and McCormick, F. (1987). A cytoplasmic protein stimulates normal N-ras p21 GTPase, but does not affect oncogenic mutants. Science *238*, 542-545.
Trahey, M., Milley, R.J., Cole, G.E., Innis, M., Paterson, H., Marshall, C.J., Hall, A., and McCormick, F. (1987). Biochemical and biological properties of the human N-ras p21 protein. Mol Cell Biol *7*, 541-544.
Ubersax, J.A., and Ferrell, J.E., Jr. (2007). Mechanisms of specificity in protein phosphorylation. Nat Rev Mol Cell Biol *8*, 530-541.
Ulintz, P.J., Yocum, A.K., Bodenmiller, B., Aebersold, R., Andrews, P.C., and Nesvizhskii, A.I. (2009). Comparison of MS(2)-only, MSA, and MS(2)/MS(3) methodologies for phosphopeptide identification. J Proteome Res *8*, 887-899.
Ulsh, L.S., and Shih, T.Y. (1984). Metabolic turnover of human c-rasH p21 protein of EJ bladder carcinoma and its normal cellular and viral homologs. Mol Cell Biol *4*, 1647-1652.
Umanoff, H., Edelmann, W., Pellicer, A., and Kucherlapati, R. (1995). The murine N-ras gene is not essential for growth and development. Proc Natl Acad Sci U S A *92*, 1709-1713.
UniProt, C. (2011). Ongoing and future developments at the Universal Protein Resource. Nucleic Acids Res *39*, D214-219.

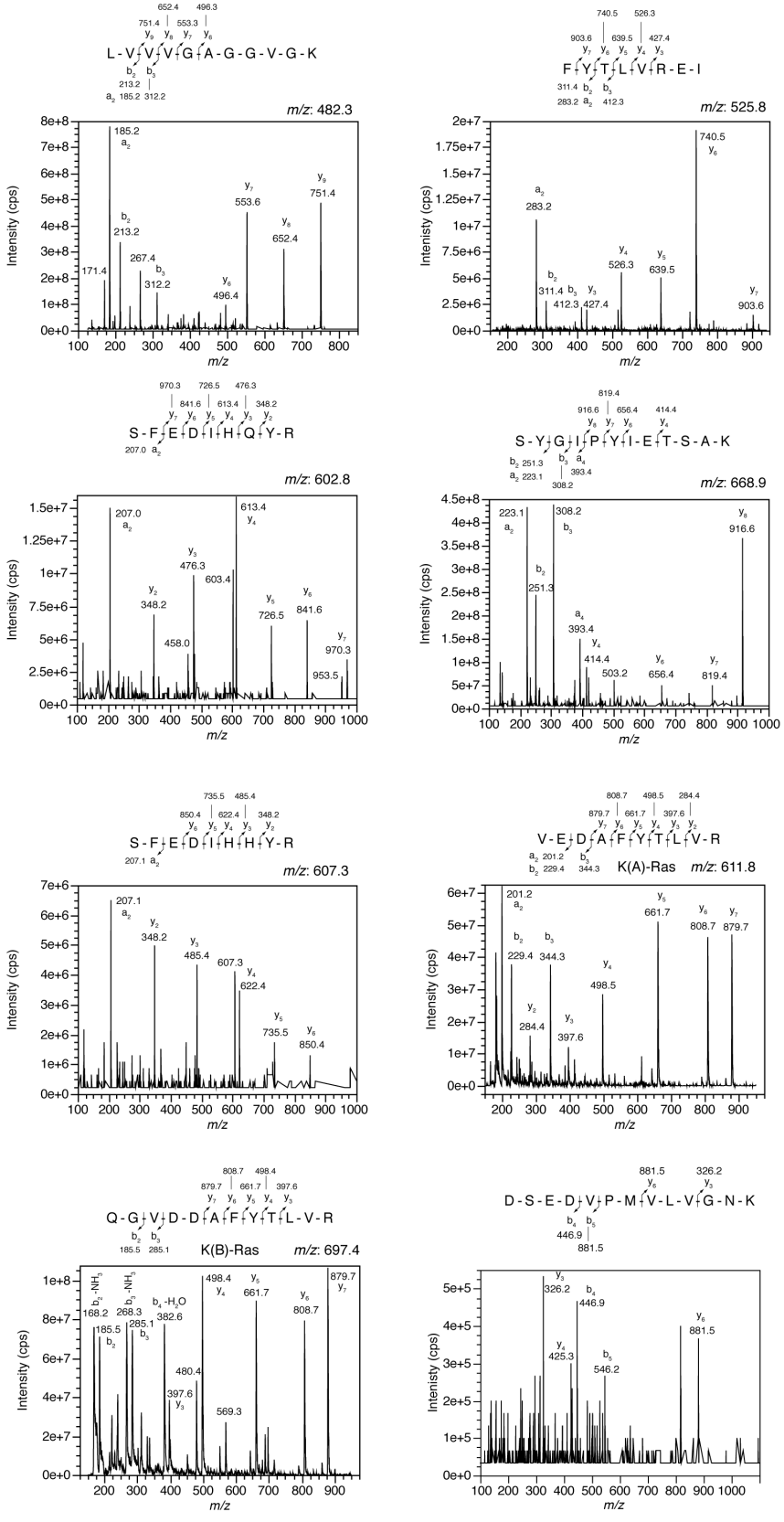
Unwin, R.D., Griffiths, J.R., Leverentz, M.K., Grallert, A., Hagan, I.M., and Whetton, A.D. (2005). Multiple reaction monitoring to identify sites of protein phosphorylation with high sensitivity. Mol Cell Proteomics *4*, 1134-1144.
Unwin, R.D., Griffiths, J.R., and Whetton, A.D. (2009). A sensitive mass spectrometric method for hypothesis-driven detection of peptide post-translational modifications: multiple reaction monitoring-initiated detection and sequencing (MIDAS). Nat Protoc *4*, 870-877.
Vaisar, T., and Urban, J. (1996). Probing the proline effect in CID of protonated peptides. Journal of Mass Spectrometry *31*, 1185-1187.
Valencia, A., Chardin, P., Wittinghofer, A., and Sander, C. (1991). The ras protein family: evolutionary tree and role of conserved amino acids. Biochemistry *30*, 4637-4648.
Valtorta, E., Misale, S., Sartore-Bianchi, A., Nagtegaal, I.D., Paraf, F., Lauricella, C., Dimartino, V., Hobor, S., Jacobs, B., Ercolani, C*., et al.* (2013). KRAS gene amplification in colorectal cancer and impact on response to EGFR-targeted therapy. Int J Cancer *133*, 1259-1265.
Vartanian, S., Bentley, C., Brauer, M.J., Li, L., Shirasawa, S., Sasazuki, T., Kim, J.S., Haverty, P., Stawiski, E., Modrusan, Z*., et al.* (2013). Identification of mutant K-Ras-dependent phenotypes using a panel of isogenic cell lines. J Biol Chem *288*, 2403-2413.
Vaughn, C.P., Zobell, S.D., Furtado, L.V., Baker, C.L., and Samowitz, W.S. (2011). Frequency of KRAS, BRAF, and NRAS mutations in colorectal cancer. Genes Chromosomes Cancer *50*, 307-312.
Vetter, I.R., and Wittinghofer, A. (2001). The guanine nucleotide-binding switch in three dimensions. Science *294*, 1299-1304.
Villalonga, P., Lopez-Alcala, C., Bosch, M., Chiloeches, A., Rocamora, N., Gil, J., Marais, R., Marshall, C.J., Bachs, O., and Agell, N. (2001). Calmodulin binds to K-Ras, but not to H- or N-Ras, and modulates its downstream signaling. Mol Cell Biol *21*, 7345-7354.
Villen, J., Beausoleil, S.A., and Gygi, S.P. (2008). Evaluation of the utility of neutral-loss-dependent MS3 strategies in large-scale phosphorylation analysis. Proteomics *8*, 4444-4452.
Villen, J., and Gygi, S.P. (2008). The SCX/IMAC enrichment approach for global phosphorylation analysis by mass spectrometry. Nat Protoc *3*, 1630-1638.
Vizan, P., Boros, L.G., Figueras, A., Capella, G., Mangues, R., Bassilian, S., Lim, S., Lee, W.N., and Cascante, M. (2005). K-ras codon-specific mutations produce distinctive metabolic phenotypes in NIH3T3 mice [corrected] fibroblasts. Cancer Res *65*, 5512-5515.
Vogelstein, B., Fearon, E.R., Hamilton, S.R., Kern, S.E., Preisinger, A.C., Leppert, M., Nakamura, Y., White, R., Smits, A.M., and Bos, J.L. (1988). Genetic alterations during colorectal-tumor development. N Engl J Med *319*, 525-532.
Voice, J.K., Klemke, R.L., Le, A., and Jackson, J.H. (1999). Four human ras homologs differ in their abilities to activate Raf-1, induce transformation, and stimulate cell motility. J Biol Chem *274*, 17164-17170.
Wada, Y., and Kadoya, M. (2003). In-gel digestion with endoproteinase LysC. J Mass Spectrom *38*, 117-118.
Walsh, A.B., and Bar-Sagi, D. (2001). Differential activation of the Rac pathway by Ha-Ras and K-Ras. J Biol Chem *276*, 15609-15615.
Wang, G., and Deschenes, R.J. (2006). Plasma membrane localization of Ras requires class C Vps proteins and functional mitochondria in Saccharomyces cerevisiae. Mol Cell Biol *26*, 3243-3255.

Wang, Q., Chaerkady, R., Wu, J., Hwang, H.J., Papadopoulos, N., Kopelovich, L., Maitra, A., Matthaei, H., Eshleman, J.R., Hruban, R.H.*, et al.* (2011). Mutant proteins as cancer-specific biomarkers. Proc Natl Acad Sci U S A *108*, 2444-2449.
Washburn, M.P., Wolters, D., and Yates, J.R., 3rd (2001). Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol *19*, 242-247.
Watson, J.T., and Sparkman, O.D. (2007). Introduction to Mass Spectrometry Instrumentation, Applications and Strategies for Data Interpretation (Chichester: John Wiley & Sons, Ltd).
Webb, C.P., Taylor, G.A., Jeffers, M., Fiscella, M., Oskarsson, M., Resau, J.H., and Vande Woude, G.F. (1998). Evidence for a role of Met-HGF/SF during Ras-mediated tumorigenesis/metastasis. Oncogene *17*, 2019-2025.
Wells, J.M., and McLuckey, S.A. (2005). Collision-induced dissociation (CID) of peptides and proteins. Methods Enzymol *402*, 148-185.
Weygant, N., Qu, D.F., Berry, W.L., May, R., Chandrakesan, P., Owen, D.B., Sureban, S.M., Ali, N., Janknecht, R., and Houchen, C.W. (2014). Small molecule kinase inhibitor LRRK2-IN-1 demonstrates potent activity against colorectal and pancreatic cancer through inhibition of doublecortin-like kinase 1. Molecular Cancer *13*.
Whitehouse, C.M., Dreyer, R.N., Yamashita, M., and Fenn, J.B. (1985). Electrospray interface for liquid chromatographs and mass spectrometers. Anal Chem *57*, 675-679.
Wilkins, M.R., Williams, K.L., Appel, R.D., and Hochstrasser, D.F. (1997). Proteome research: new frontiers in functional genomics (Springer-Verlag).
Willingham, M.C., Pastan, I., Shih, T.Y., and Scolnick, E.M. (1980). Localization of the src gene product of the Harvey strain of MSV to plasma membrane of transformed cells by electron microscopic immunocytochemistry. Cell *19*, 1005-1014.
Willumsen, B.M., Norris, K., Papageorge, A.G., Hubbert, N.L., and Lowy, D.R. (1984). Harvey murine sarcoma virus p21 ras protein: biological and biochemical significance of the cysteine nearest the carboxy terminus. EMBO J *3*, 2581-2585.
Wilm, M., and Mann, M. (1996). Analytical properties of the nanoelectrospray ion source. Anal Chem *68*, 1-8.
Wilm, M., Shevchenko, A., Houthaeve, T., Breit, S., Schweigerer, L., Fotsis, T., and Mann, M. (1996). Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. Nature *379*, 466-469.
Wilm, M.S., and Mann, M. (1994). Electrospray and Taylor-Cone theory, Dole's beam of macromolecules at last? International Journal of Mass Spectrometry and Ion Processes *136*, 167-180.
Wisniewski, J.R., Zougman, A., Nagaraj, N., and Mann, M. (2009). Universal sample preparation method for proteome analysis. Nat Methods *6*, 359-362.
Wold, F. (1981). In vivo chemical modification of proteins (post-translational modification). Annu Rev Biochem *50*, 783-814.
Wolff, M.M., and Stephens, W.E. (1953). A Pulsed Mass Spectrometer with Time Dispersion. Review of Scientific Instruments *24*, 616-617.
Wolters, D.A., Washburn, M.P., and Yates, J.R., 3rd (2001). An automated multidimensional protein identification technology for shotgun proteomics. Anal Chem *73*, 5683-5690.
Yan, J., Roy, S., Apolloni, A., Lane, A., and Hancock, J.F. (1998). Ras isoforms vary in their ability to activate Raf-1 and phosphoinositide 3-kinase. J Biol Chem *273*, 24052-24056.

Yao, X., Freas, A., Ramirez, J., Demirev, P.A., and Fenselau, C. (2001). Proteolytic 18O labeling for comparative proteomics: model studies with two serotypes of adenovirus. Anal Chem *73*, 2836-2842.
Ye, M., Shima, F., Muraoka, S., Liao, J.L., Okamoto, H., Yamamoto, M., Tamura, A., Yagi, N., Ueki, T., and Kataoka, T. (2005). Crystal structure of M-Ras reveals a GTP-bound "off" state conformation of Ras family small GTPases. Journal of Biological Chemistry *280*, 31267-31275.
Yeung, T., Terebiznik, M., Yu, L., Silvius, J., Abidi, W.M., Philips, M., Levine, T., Kapus, A., and Grinstein, S. (2006). Receptor activation alters inner surface potential during phagocytosis. Science *313*, 347-351.
Yost, R.A., and Enke, C.G. (1979). Triple quadrupole mass spectrometry for direct mixture analysis and structure elucidation. Anal Chem *51*, 1251-1264.
Young, A., Lou, D., and McCormick, F. (2013). Oncogenic and wild-type Ras play divergent roles in the regulation of mitogen-activated protein kinase signaling. Cancer Discov *3*, 112-123.
Yu, W., Vath, J.E., Huberty, M.C., and Martin, S.A. (1993). Identification of the facile gas-phase cleavage of the Asp-Pro and Asp-Xxx peptide bonds in matrix-assisted laser desorption time-of-flight mass spectrometry. Anal Chem *65*, 3015-3023.
Zakett, D., Flynn, R.G.A., and Cooks, R.G. (1978). Chlorine Isotope Effects in Mass-Spectrometry by Multiple Reaction Monitoring. Journal of Physical Chemistry *82*, 2359-2362.
Zubarev, R.A., Kelleher, N.L., and McLafferty, F.W. (1998). Electron capture dissociation of multiply charged protein cations. A nonergodic process. Journal of the American Chemical Society *120*, 3265-3266.

6 – SUPPLEMENTARY MATERIAL

6.1 PROTEOTYPIC RAS PEPTIDE MS/MS SPECTRA



Continued onto next page

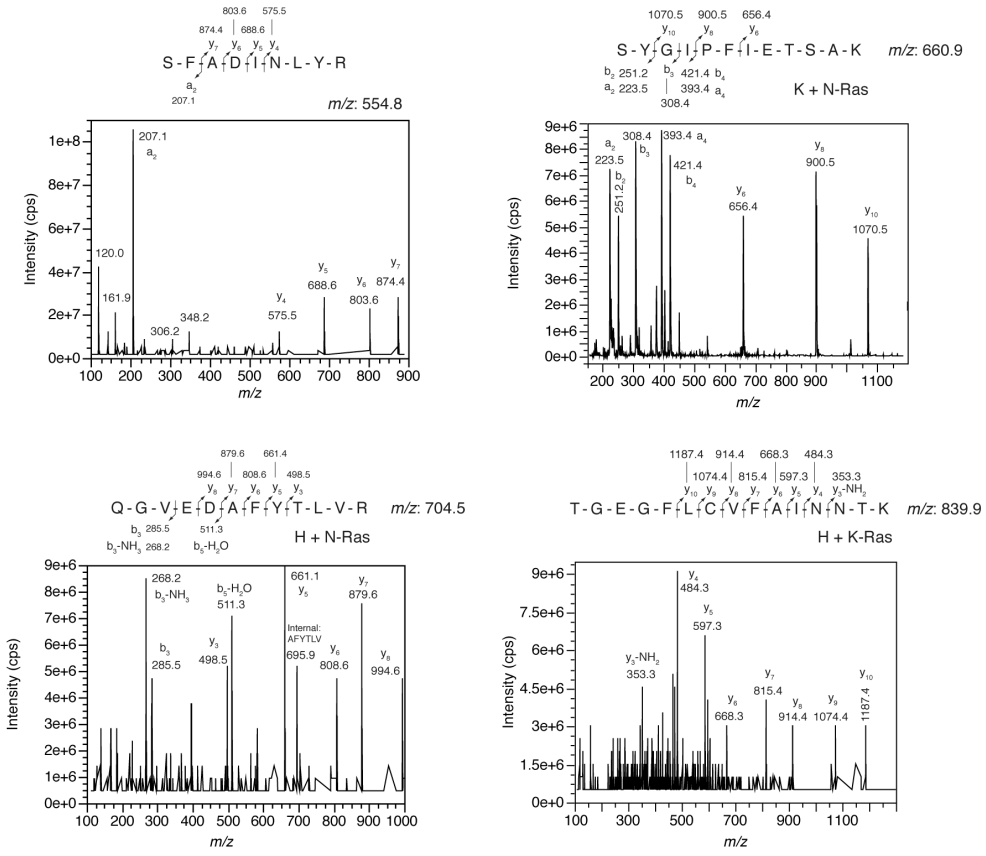


Fig. 6.1. MS/MS spectra of the Ras-specific proteotypic peptides identified in the present thesis. All spectra were acquired on using a 4000 QTRAP using isotope-labelled Ras peptides (Lys8, Arg10).

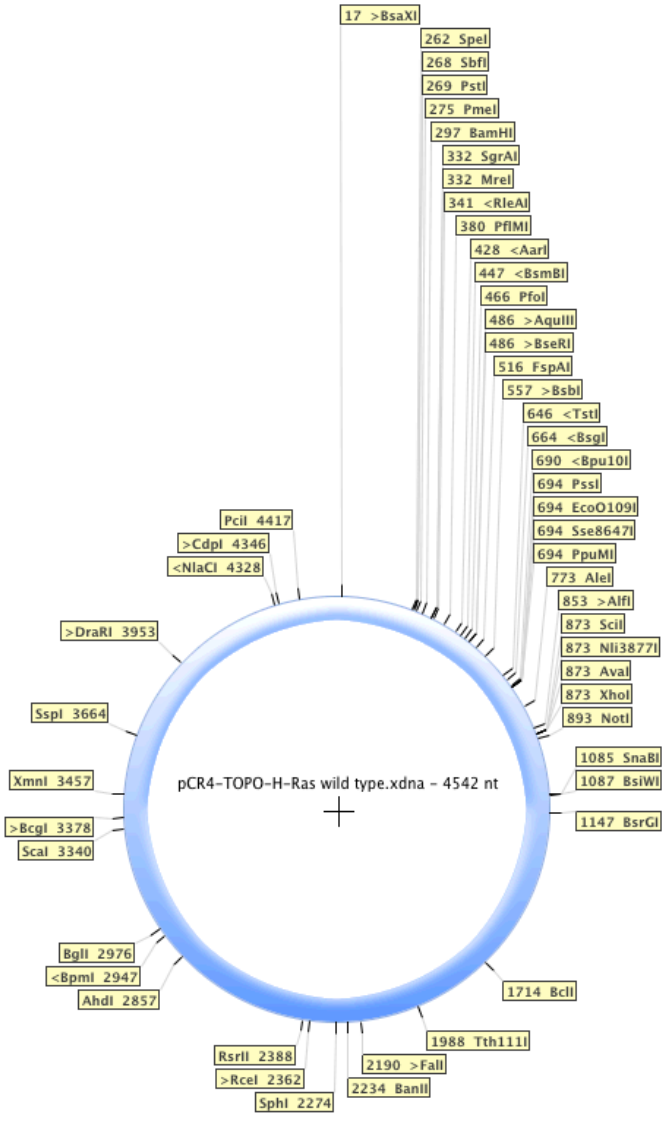
6.2 CELL COUNTING AND LYSATE VOLUME MEASUREMENTS FOR RAS PSAQ ANALYSIS

RAS mutation status	Replicate	Cells per square on Haemocytometer (0.1 μL)						Average	SD	CV(%)	Volume collected (mL)	Total number of cell	Lysate volume (μL)	Concentration (μg/μL)	Volume of lysate used (μL)	% lysate used	Number of cells used
		1	2	3	4	5	6										
KRAS G12A	1	152	147	150	135	143	104	138.50	17.94	12.95	9	12,465,000	204	4.40	22.74	11.15	1,389,274
KRAS G12A	2	112	134	127	118	108	96	115.83	13.63	11.77	9	10,425,000	190	4.74	21.10	11.10	1,157,670
KRAS G12A	3	64	107	91	94	124	124	100.67	22.85	22.70	9	9,060,000	185	3.74	26.71	14.44	1,308,307
KRAS G12C	1	125	148	117	133	138	111	128.67	13.72	10.66	9	11,580,000	209	7.17	13.95	6.67	772,697
KRAS G12C	2	97	97	131	109	113	126	112.17	14.26	12.71	9	10,095,000	190	5.23	19.13	10.07	1,016,614
KRAS G12C	3	91	95	111	98	94	97	97.67	6.98	7.14	9	8,790,000	178	6.07	16.48	9.26	813,571
KRAS G12D	1	122	145	118	168	145	152	141.67	18.81	13.28	9	12,750,000	187	6.25	16.01	8.56	1,091,681
KRAS G12D	2	124	142	123	148	153	146	139.33	12.77	9.16	9	12,540,000	183	5.02	19.91	10.88	1,364,488
KRAS G12D	3	123	129	92	97	103	84	104.67	17.76	16.97	9	9,420,000	182	5.65	17.69	9.72	915,676
KRAS G12R	1	145	186	163	145	152	134	154.17	18.28	11.86	9	13,875,000	167	5.70	17.53	10.50	1,456,739
KRAS G12R	2	197	155	141	167	185	173	169.67	20.19	11.90	9	15,270,000	173	6.28	15.93	9.21	1,405,799
KRAS G12R	3	112	120	100	99	111	109	108.50	7.92	7.30	9	9,765,000	180	4.35	22.97	12.76	1,246,315
KRAS G12S	1	146	126	143	136	114	117	130.33	13.43	10.30	9	11,730,000	191	4.02	24.90	13.03	1,528,970
KRAS G12S	2	118	114	104	102	134	128	116.67	12.75	10.93	9	10,500,000	170	3.68	27.19	16.00	1,679,583
KRAS G12S	3	101	84	90	87	83	62	84.50	12.79	15.13	9	7,605,000	173	3.34	29.97	17.33	1,317,596
KRAS G12V	1	167	178	167	171	221	159	177.17	22.35	12.61	9	15,945,000	210	6.44	15.54	7.40	1,179,685
KRAS G12V	2	155	134	144	142	162	158	149.17	10.82	7.25	9	13,425,000	188	6.10	16.38	8.72	1,170,009
KRAS G12V	3	129	152	155	152	113	169	145.00	20.27	13.98	9	13,050,000	190	6.16	16.22	8.54	1,114,338
KRAS G13D	1	87	98	92	72	62	-	82.20	14.84	18.05	9	7,398,000	128	4.30	23.25	18.16	1,343,596
KRAS G13D	2	45	56	43	37	33	40	42.33	7.94	18.76	9	3,810,000	140	4.36	22.92	16.37	623,724
KRAS G13D	3	77	113	139	109	93	110	106.83	20.83	19.49	9	9,615,000	123	3.32	30.15	24.51	2,356,640
HRAS G12V	1	64	75	50	75	41	53	59.67	13.97	23.41	9	5,370,000	187	4.80	20.85	11.15	598,763
HRAS G12V	2	62	61	73	45	51	67	59.83	10.28	17.19	9	5,385,000	180	3.22	31.09	17.27	930,059
HRAS G12V	3	46	57	52	49	52	64	53.33	6.38	11.96	9	4,800,000	175	3.84	26.06	14.89	714,820
NRAS G12V	1	141	134	93	135	126	157	131.00	21.31	16.27	9	11,790,000	190	6.46	15.49	8.15	961,002
NRAS G12V	2	130	111	123	123	95	124	117.67	12.71	10.80	9	10,590,000	190	4.13	24.20	12.74	1,348,924
NRAS G12V	3	102	113	145	108	101	109	113.00	16.31	14.43	9	10,170,000	188	5.35	18.69	9.94	1,011,153
WT	1	76	97	112	104	109	98	99.33	12.86	12.95	9	8,940,000	223	3.61	27.74	12.44	1,112,052
WT	2	101	101	114	122	125	125	114.67	11.33	9.88	9	10,320,000	199	4.76	21.01	10.56	1,089,774
WT	3	132	128	140	145	110	137	132.00	12.31	9.33	9	11,880,000	187	6.13	16.31	8.72	1,035,966

Table 6.1. Cell counting and lysate volume measurements in the Ras PSAQ analysis. Ras mutational status indicates which heterozygous knock-in Ras mutation was present.

6.3 PLASMID MAPS

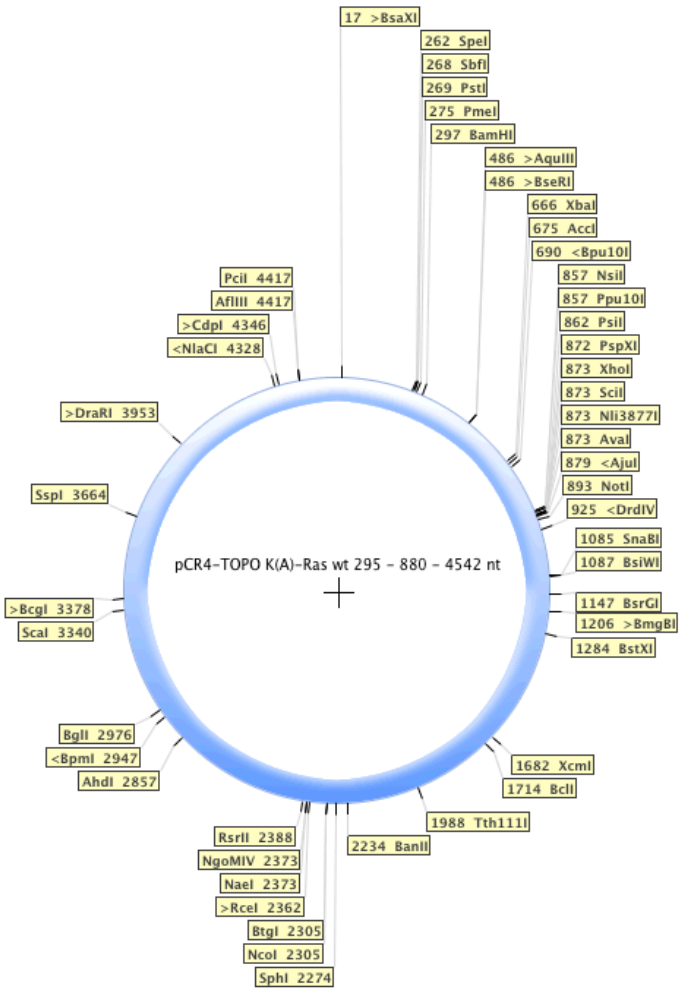
Name of construct: pCR4-TOPO-H-Ras wild-type
Origin of vector: pCR4-TOPO kit
Origin of insert: PCR of GFP-H-Ras.
Vector size: 3956 bp
Insert size: 570 bp
ORF starts: 303 – 872 bp
Constructed by: Craig Mageean (Univ. Liverpool)
Resistance: Kanamycin/ ampicillin
Date: 06.05.11
Maxiprep conc: 0.22 ug/ uL
Comments: Fully sequenced H-Ras wild-type ORF with no mutations. Contains ATG and stop codon with added BamH1 and Xho1 restriction sites to cut out the entire sequence.



Alignment of Ras sequence from pCR4-TOPO-H-Ras with H-Ras ORF (taken from Uniprot #P0112):

```
pCR4-TOPO-H-RasG12      CCATGACGGAATATAAGCTGGTGGTGGTGGGCGCCGGCGGTGTGGGCAAGAGTGCGCTGA 360
H-Ras_ORF               --ATGACGGAATATAAGCTGGTGGTGGTGGGCGCCGGCGGTGTGGGCAAGAGTGCGCTGA 58
                          ************************************************************

pCR4-TOPO-H-RasG12      CCATCCAGCTGATCCAGAACCATTTTGTGGACGAATACGACCCCACTATAGAGGATTCCT 420
H-Ras_ORF               CCATCCAGCTGATCCAGAACCACTTTGTGGACGAATACGACCCCACTATAGAGGATTCCT 118
                        ********************* ************************************

pCR4-TOPO-H-RasG12      ACCGGAAGCAGGTGGTCATTGATGGGGAGACGTGCCTGTTGGACATCCTGGATACCGCCG 480
H-Ras_ORF               ACCGGAAGCAGGTGGTCATTGATGGGGAGACGTGCCTGTTGGACATCCTGGATACCGCCG 178
                        ************************************************************

pCR4-TOPO-H-RasG12      GCCAGGAGGAGTACAGCGCCATGCGGGACCAGTACATGCGCACCGGGGAGGGCTTCCTGT 540
H-Ras_ORF               GCCAGGAGGAGTACAGCGCCATGCGGGACCAGTACATGCGCACCGGGGAGGGCTTCCTGT 238
                        ************************************************************

pCR4-TOPO-H-RasG12      GTGTGTTTGCCATCAACAACACCAAGTCTTTTGAGGACATCCACCAGTACAGGGAGCAGA 600
H-Ras_ORF               GTGTGTTTGCCATCAACAACACCAAGTCTTTTGAGGACATCCACCAGTACAGGGAGCAGA 298
                        ************************************************************

pCR4-TOPO-H-RasG12      TCAAACGGGTGAAGGACTCGGATGACGTGCCCATGGTGCTGGTGGGGAACAAGTGTGACC 660
H-Ras_ORF               TCAAACGGGTGAAGGACTCGGATGACGTGCCCATGGTGCTGGTGGGGAACAAGTGTGACC 358
                        ************************************************************

pCR4-TOPO-H-RasG12      TGGCTGCACGCACTGTGGAATCTCGGCAGGCTCAGGACCTCGCCCGAAGCTACGGCATCC 720
H-Ras_ORF               TGGCTGCACGCACTGTGGAATCTCGGCAGGCTCAGGACCTCGCCCGAAGCTACGGCATCC 418
                        ************************************************************

pCR4-TOPO-H-RasG12      CCTACATCGAGACCTCGGCCAAGACCCGGCAGGGAGTGGAGGATGCCTTCTACACGTTGG 780
H-Ras_ORF               CCTACATCGAGACCTCGGCCAAGACCCGGCAGGGAGTGGAGGATGCCTTCTACACGTTGG 478
                        ************************************************************

pCR4-TOPO-H-RasG12      TGCGTGAGATCCGGCAGCACAAGCTGCGGAAGCTGAACCCTCCTGATGAGAGTGGCCCCG 840
H-Ras_ORF               TGCGTGAGATCCGGCAGCACAAGCTGCGGAAGCTGAACCCTCCTGATGAGAGTGGCCCCG 538
                        ************************************************************

pCR4-TOPO-H-RasG12      GCTGCATGAGCTGCAAGTGTGTGCTCTCCTGACTCGAGACAAGGGCGAATTCGCGGCCGC 900
H-Ras_ORF               GCTGCATGAGCTGCAAGTGTGTGCTCTCC------------------------------- 567
                        ****************************
```

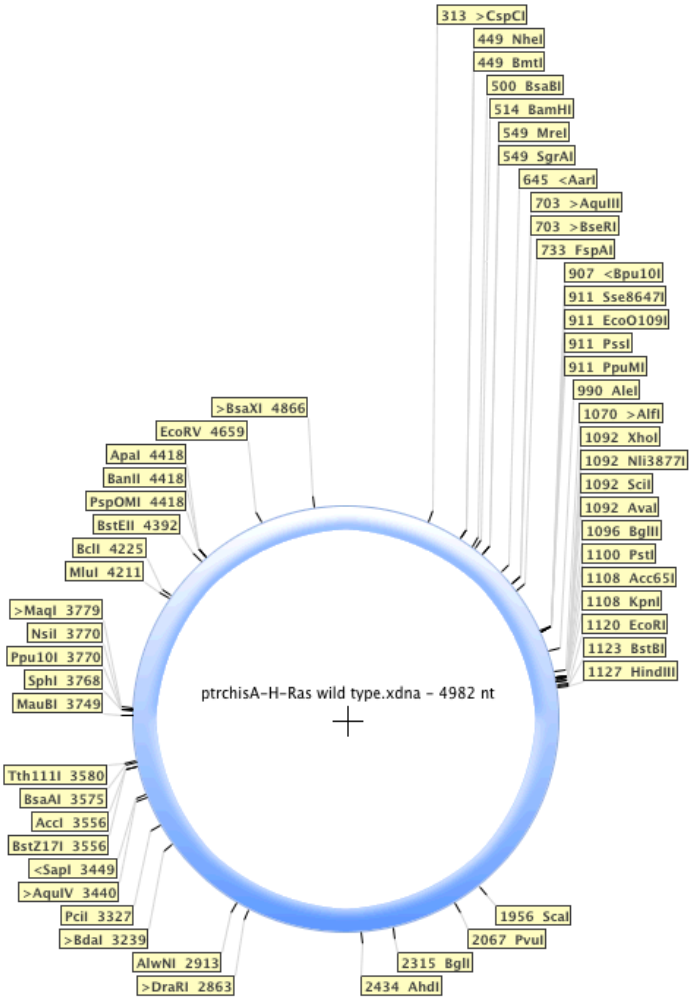
Name of construct: pCR4-TOPO-K(A)-Ras wild-type
Origin of vector: pCR4-TOPO kit
Origin of insert: PCR of GFP-K(A)-Ras
Vector size: 3956 bp
Insert size: 585 bp
ORF starts: 295 – 880 bp
Constructed by: Craig Mageean (Univ. Liverpool)
Resistance: Kanamycin/ ampicillin
Date: 12.04.11
Maxiprep conc: 0.14 ug/ uL
Comments: Fully sequenced K(A)-Ras wild-type ORF with no mutations. Contains ATG and stop codon with added BamH1 and Xho1 restriction sites to cut out the entire sequence.



Alignment of Ras sequence from pCR4-TOPO-K(A)-Ras with K(A)-Ras ORF (taken from Uniprot #P0116).

```
pCR4-TOPO-K_A_-RasG12      GGATCCATGACTGAATATAAACTTGTGGTAGTTGGAGCTGGTGGCGTAGG 100
K_A_-Ras_ORF               ------ATGACTGAATATAAACTTGTGGTAGTTGGAGCTGGTGGCGTAGG 44
                                 ******************************************

pCR4-TOPO-K_A_-RasG12      CAAGAGTGCCTTGACGATACAGCTAATTCAGAATCATTTTGTGGACGAAT 150
K_A_-Ras_ORF               CAAGAGTGCCTTGACGATACAGCTAATTCAGAATCATTTTGTGGACGAAT 94
                           **************************************************

pCR4-TOPO-K_A_-RasG12      ATGATCCAACAATAGAGGATTCCTACAGGAAGCAAGTAGTAATTGATGGA 200
K_A_-Ras_ORF               ATGATCCAACAATAGAGGATTCCTACAGGAAGCAAGTAGTAATTGATGGA 144
                           **************************************************

pCR4-TOPO-K_A_-RasG12      GAAACCTGTCTCTTGGATATTCTCGACACAGCAGGTCAAGAGGAGTACAG 250
K_A_-Ras_ORF               GAAACCTGTCTCTTGGATATTCTCGACACAGCAGGTCAAGAGGAGTACAG 194
                           **************************************************

pCR4-TOPO-K_A_-RasG12      TGCAATGAGGGACCAGTACATGAGGACTGGGGAGGGCTTTCTTTGTGTAT 300
K_A_-Ras_ORF               TGCAATGAGGGACCAGTACATGAGGACTGGGGAGGGCTTTCTTTGTGTAT 244
                           **************************************************

pCR4-TOPO-K_A_-RasG12      TTGCCATAAATAATACTAAATCATTTGAAGATATTCACCATTATAGAGAA 350
K_A_-Ras_ORF               TTGCCATAAATAATACTAAATCATTTGAAGATATTCACCATTATAGAGAA 294
                           **************************************************

pCR4-TOPO-K_A_-RasG12      CAAATTAAAAGAGTTAAGGACTCTGAAGATGTACCTATGGTCCTAGTAGG 400
K_A_-Ras_ORF               CAAATTAAAAGAGTTAAGGACTCTGAAGATGTACCTATGGTCCTAGTAGG 344
                           **************************************************

pCR4-TOPO-K_A_-RasG12      AAATAAATGTGATTTGCCTTCTAGAACAGTAGACACAAAACAGGCTCAGG 450
K_A_-Ras_ORF               AAATAAATGTGATTTGCCTTCTAGAACAGTAGACACAAAACAGGCTCAGG 394
                           **************************************************

pCR4-TOPO-K_A_-RasG12      ACTTAGCAAGAAGTTATGGAATTCCTTTTATTGAAACATCAGCAAAGACA 500
K_A_-Ras_ORF               ACTTAGCAAGAAGTTATGGAATTCCTTTTATTGAAACATCAGCAAAGACA 444
                           **************************************************

pCR4-TOPO-K_A_-RasG12      AGACAGAGAGTGGAGGATGCTTTTTATACATTGGTGAGAGAGATCCGACA 550
K_A_-Ras_ORF               AGACAGAGAGTGGAGGATGCTTTTTATACATTGGTGAGGGAGATCCGACA 494
                           ************************************ ***********

pCR4-TOPO-K_A_-RasG12      ATACAGATTGAAAAAAATCAGCAAAGAAGAAAAGACTCCTGGCTGTGTGA 600
K_A_-Ras_ORF               ATACAGATTGAAAAAAATCAGCAAAGAAGAAAAGACTCCTGGCTGTGTGA 544
                           **************************************************

pCR4-TOPO-K_A_-RasG12      AAATTAAAAAATGCATTATAATGTAACTCGAGCCAAGGGCGAATTCGTTT 650
K_A_-Ras_ORF               AAATTAAAAAATGCATTATAATGTAA---------------------- 570
                           *************************
```

Name of construct: ptrchisA-H-Ras wild-type
Origin of insert: Sequence cut from pCR4-TOPO-H-Ras with BamH1 and Xho1
Vector size: 4414 bp
Insert size: 570 bp
ORF starts: 515 – 1089 bp
Constructed by: Craig Mageean (Univ. Liverpool)
Resistance: Ampicillin
Date: 18.05.11
Maxiprep conc: 41.1 ng/ uL
Comments: Contains ATG and stop codon with added BamH1 and Xho1 restriction sites to cut out the entire sequence. Exact same sequence as PCR4-TOPO-H-Ras wild-type.



Name of construct: ptrchisA-K(A)-Ras wild-type
Origin of insert: Sequence cut from pCR4-TOPO-H-Ras with BamH1 and Xho1
Vector size: 4414 bp
Insert size: 585 bp
ORF starts: 413 – 1090 bp
Constructed by: Craig Mageean (Univ. Liverpool)
Resistance: Ampicillin
Date: 18.05.11
Maxiprep conc: 38 ng/ uL
Comments: Contains ATG and stop codon with added BamH1 and Xho1 restriction sites to cut out the entire sequence. Exact same sequence as PCR4-TOPO-K(A)-Ras wild-type.